ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 51

FASC. 1—2

SZEGED, 1987

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

51. KÖTET

FASC. 1—2

SZEGED, 1987

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 51

SZEGED, 1987

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

51. KÖTET

SZEGED, 1987

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

Acta Sci. Math., 51 (1987), 3—15

The finitely based varieties of graph algebras

KIRBY A. BAKER*, GEORGE F. MCNULTY** and HEINRICH WERNER

Dedicated to the memory of András P. Huhn

1. Introduction

SHALLON [17] proposed a method of making graphs into algebras (algebraic systems) for which even a small finite graph can have a rich theory of equations with unusual properties.

Specifically, for any graph G (possibly with loops at the vertices but without multiple edges), add one new element ∞ to obtain the set $G^{\#}=G\cup\{\infty\}$, and define a binary operation $*$ on $G^{\#}$ by $x*y=x$ if x and y are joined by an edge, and $x*y=\infty$ otherwise. The *Shallon graph algebra* is the pair $\mathbf{G}^{\#}=\langle G^{\#};*\rangle$. Such algebras have been investigated in [9], [12], [14], [15], [16] and [17].

An *equational basis* for an algebra is a list of equations, true in the algebra, of which all equations true in the algebra are logical consequences. LYNDON [7] discovered the surprising fact that a finite algebra may have no finite equational basis. His example had seven elements and one binary operation. MURSKIĬ [10] later found a three-element example. Such algebras are said to be *nonfinitely based.*

SHALLON [17] showed that for the looped graph L_3 of Figure 1, $\mathbf{L}_3^{\#}$ is nonfinitely based. She also noted that Murskiĭ's example is $\mathbf{M}^{\#}$, and gave additional examples.

In a further development, PERKINS [13] and MURSKIĬ [11] discovered that some algebras, including Murskiĭ's example, are nonfinitely based in a contagious way: If the algebra in question is a subalgebra or homomorphic image of another finite algebra, then that algebra too is nonfinitely based. More generally, an algebra A is said to be *inherently nonfinitely based* [13] if A is contained in some locally finite

Received June 20, 1986, and in revised form November 14, 1986.

* Supported in part by NSF Grant # MCS 81-02519.

** Supported in part by NSF Grant # ISP-11451, a Fulbright-Hays Grant from the Phillipine-American Educational Foundation, and a Fellowship from the Alexander von Humboldt Stiftung.

variety but in no finitely based locally finite variety. (A variety is said to be locally finite if its finitely generated members are finite.) In [1], the authors showed that in fact all four graphs of Figure 1 have inherently nonfinitely based graph algebras.

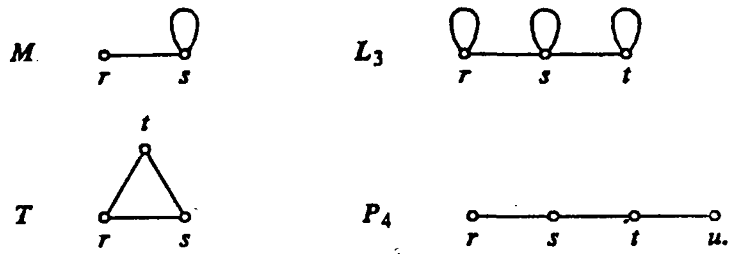


Figure 1

It follows that any graph with one of these four as an induced subgraph also has an inherently nonfinitely based graph algebra.

We obtain the following facts.

1.1. Theorem. *A graph G has a finitely based graph algebra if and only if G has no [induced] subgraph isomorphic to one of the four graphs of Figure* 1.

1.2. Corollary. *If a graph algebra is not finitely based, then it is inherently nonfinitely based.*

Indeed, in Section 2 it is shown that the graph algebras of graphs not containing one of the four graphs of Figure 1 are members of a specific variety (Proposition 2.4), and in Section 3 it is shown that all graph algebras in that variety are finitely based (Theorem 3.1). These facts, together with the result quoted from [1], constitute a proof of Theorem 1.1 and Corollary 1.2.

Further, we show that for the specific variety just mentioned, all subvarieties are finitely based. For each of these we give explicit defining equations. The lattice of subvarieties is discussed in Sections 4, 5.

Graph algebras are natural candidates for applying the methods of [1]. They are locally finite and have absorbing elements. Further, it is not hard to show that the variety generated by a class of graph algebras must be locally finite and generated by a single graph algebra. Some interesting algebraic features, such as simplicity and subdirect irreducibility, can be easily discerned by inspection of the graph.

It simplifies the arguments below to consider *augmented graph algebras*. For a graph G, the corresponding augmented graph algebra, here denoted by G^*, is obtained from $G^\#$ by declaring the absorbing element ∞ to be a nullary operation (distinguished constant). We actually prove Theorem 1.1 for the case of augmented graph algebras and then in Section 5 explain the modifications necessary for the unaugmented case.

We denote by $\langle G\rangle$ the variety generated by the augmented graph algebra G^*. For each $k=1, 2, \ldots,$ L_k and P_k will denote k-vertex graphs in the form of a path, with and without loops, respectively, as in the diagrams of P_4 and L_3 in Figure 1. In particular, L_1 is the graph with a single looped vertex and P_1 is the graph with a single unlooped vertex. For graphs G and H, $G+H$ will denote the disjoint union of G and H, with no edges between the two.

In most respects we follow the terminology and notation of [1] and [2]. Additional valuable references are [3] and [8]. We use the notation G both for a graph and for its vertex set. By a subgraph we always mean an induced subgraph.

The authors are grateful to the referee for detailed suggestions.

2. A characterization of graphs with excluded subgraphs

By a *complete* graph we mean a graph in which every two vertices are joined by an edge and in which there is a loop at each vertex. A graph G is said to be *bipartite-complete* if G decomposes into two disjoint subsets, $G=G_0+G_1$, and there is an edge between every member of G_0 and every member of G_1 but no other edges; in particular, there are no loops.

2.1. Proposition. *For a graph* G, *the following conditions are equivalent*:
(a) *G has no subgraph isomorphic to* M, T, P_4, *or* L_3;
(b) *each connected component of* G *is complete or bipartite-complete.*

Proof. Trivially, (b)$\Rightarrow$(a). For (a)$\Rightarrow$(b): Let G be a connected graph that does not contain M, T, P_4, or L_3 as a subgraph. Since M is not a subgraph of G, G has either no loops at all or a loop at every vertex.

Case 1: All vertices of G have loops. Since L_3 is not a subgraph of G, any two vertices are connected by an edge and hence G is complete.

Case 2: No vertex of G has a loop. Since T and P_4 are not subgraphs of G, each path of three edges must have an extra edge between its beginning and end vertices, as portrayed in Figure 2.

Figure 2

Thus every vertex has an edge to any other vertex at an odd distance, but no edge to any vertex at an even distance. Therefore G is bipartite-complete.

We shall now need more information about the class of graphs G whose augmented graph algebras G^* belong to a given variety **V**. A *groupoid with absorbing* *element* is an algebra with one binary operation and a nullary operation ∞ that is absorbing for the binary operation.

2.2. Lemma. *Let* $\mathbf{V}$ *be a variety of groupoids with absorbing element* ∞. *The class of graphs* G *with* $G^* \in \mathbf{V}$ *is closed under formation of*

(i) *subgraphs;*
(ii) *strong homomorphic images;*
(iii) *Cartesian products;*
(iv) *disjoint (i.e., disconnected) unions.*

(For a converse see PÖSCHEL and WESSEL [14] and Kiss [4]; for related results on digraph algebras see PÖSCHEL [14], [15].)

Proof. For (i): If H is a subgraph of G then clearly H^* is a subalgebra of G^*.

For (ii): Let $F: G \to H$ be a strong homomorphism of G onto H. In other words, $f(x)$ is adjacent to $f(y)$ if and only if x is adjacent to y. Extend f to G^* by setting $f(\infty) = \infty$. Then f becomes a homomorphism of G^* onto H^*.

For (iii): The subset $B = \{x \in \prod_{i \in I} G_i^* \mid x(i) = \infty$ for some $i\}$ defines a congruence $\Theta = \mathrm{id} \cup (B \times B)$ on $\prod_{i \in I} G_i^*$, and $(\prod_{i \in I} G_i)^* \cong (\prod_{i \in I} G_i^*)/\Theta$.

For (iv): The subset $C = \{x \in \prod_{i \in I} G_i^* \mid x(i) \neq \infty$ for at most one $i\}$ defines a subalgebra of $\prod_{i \in I} G_i^*$ isomorphic to $(\sum_{i \in I} G_i)^*$.

This lemma enables the construction of many augmented graph algebras in a variety $\mathbf{V}$ containing G^* for some graph G.

2.3. Proposition. *Suppose* $\mathbf{V}$ *is a variety containing* G^* *for some graph* G.

(a) *If* G *contains a connected component that is complete and has at least two vertices, then* $\mathbf{V}$ *contains all graphs whose connected components are complete.*

(b) *If* G *contains a connected component that is bipartite-complete and has at least three vertices, then* $\mathbf{V}$ *contains all graphs whose connected components are bipartite-complete.*

Proof. By 2.2-(iv), it suffices to prove that $\mathbf{V}$ contains all complete graphs in case (a) and all bipartite-complete graphs in case (b).

For (a): By (i) $\mathbf{V}$ contains L_2^*. Every complete graph is a subgraph of some power of L_2 and so yields an augmented graph algebra in $\mathbf{V}$.

For (b): If G contains a connected component that is bipartite-complete and has at least three vertices, then by (i) $\mathbf{V}$ contains P_3^*. $P_3 \times P_3$ has the two components X and Q of Figure 3, and every bipartite-complete graph is a subgraph of some power of Q; hence 2.2-(iii) and 2.2-(i) apply to show that $\mathbf{V}$ contains every bipartite-complete graph.

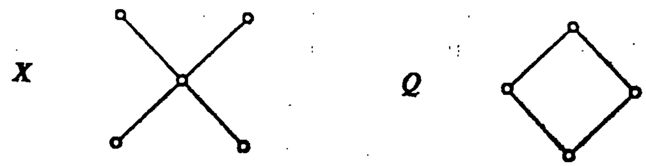


Figure 3

In contrast to Proposition 2.3, these graphs G have the property that every power of G has only copies of G itself as connected components: L_1, P_1, and P_2.

2.4. Proposition. *All augmented graph algebras that are not inherently nonfinitely based belong to the variety* $\langle P_2+L_2\rangle$.

(The converse is true and forms part of Theorem 3.2.)

Proof. The variety $\mathbf{V}=\langle P_2+L_2\rangle$ contains L_2^* and therefore contains all complete graphs by Proposition 2.3. By 2.2-(i) and 2.2-(iii), $\mathbf{V}$ contains $(P_2\times L_2)^*$. But $P_2\times L_2\cong Q$, a bipartite-complete graph of more than two elements, so that $\mathbf{V}$ contains all bipartite-complete graphs by Proposition 2.3. Now, any augmented graph algebra that is not inherently nonfinitely based has components of only these two kinds, by Lemma 2.1; and so is in $\mathbf{V}$ by 2.2-(iv).

3. Equations and the finitely based varieties

Since we want to give finite equational bases for the varieties in question, we must examine the evaluations of (groupoid) terms in augmented graph algebras.

Because graph varieties have an absorbing element ∞, their equations have a particular form:

3.1. Lemma. *Let* $\mathbf{V}$ *be a variety with absorbing element* ∞ *and* $\sigma=\tau$ *an equation true in* $\mathbf{V}$. *Then either* $\sigma=\tau$ *is a regular equation (i.e., in* σ *and* τ *the same variables occur) or else the equations* $\sigma=\infty$ *and* $\tau=\infty$ *also hold in* $\mathbf{V}$.

Proof. Assume some variable x occurs in σ but not in τ. Replace x by ∞ and leave all other variables unchanged. Then σ evaluates to ∞ and hence $\tau=\infty$ holds in $\mathbf{V}$. Since $\sigma=\tau$ also holds in $\mathbf{V}$, $\sigma=\infty$ follows.

A term τ takes the value of its leftmost variable or ∞, depending on whether or not for each subterm $\sigma_2\cdot\sigma_3$ of τ the values of the leftmost variables of σ_2 and σ_3 are connected by an edge in the underlying graph or not.

Here are some simple examples of equations true in every augmented graph algebra; see Figure 4 for illustrations of substitutions that give values other than ∞:

$$x\infty = \infty = \infty x, \tag{0}$$

$$xy = (xy)y, \tag{1}$$

$$x(yz) = (xy)(yz), \tag{2}$$

$$(xy)z = (xz)y, \tag{3}$$

$$xy = x(yx), \tag{4}$$

$$x((yz)u) = (x(yz))(yu). \tag{5}$$

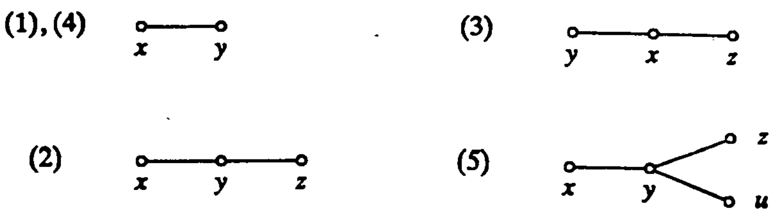


Figure 4

If all vertices of a graph are looped (as in a complete graph) its augmented graph algebra satisfies the *idempotent law*

$$xx = x. \tag{id}$$

However, if no vertex has a loop (as in a bipartite-complete graph) its augmented graph algebra satisfies the *nilpotent law*

$$xx = yy. \tag{np}$$

These two laws are contradictory, in the sense that they together imply $x=y$, the equation of the trivial variety.

Additional equations true in P_2^* and L_2^* are these; illustrations giving values other than ∞ are shown in Figure 5.

$$x(y(zu)) = (x(yz))(uz), \tag{6}$$

$$(x(yz))(uv) = (x(yv))(uz), \tag{7}$$

$$x(xy) = x(yy), \tag{8}$$

$$(xx)(yz) = (x(yy))(zz). \tag{9}$$

(The proofs are omitted. The solid edges in the graph diagrams are edges that must exist in order that the terms have values given by their leftmost variables.)

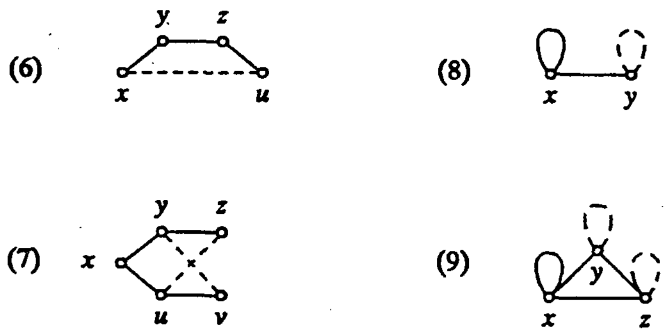


Figure 5

Another useful consequence of these equations is that

(10) $$x(yy) \underset{(1)}{=} x(yy)(yy) \underset{(9)}{=} (xx)(yy) \underset{(4)}{=} (x(xx))(yy).$$

The difficulty with varieties generated by graph algebras, augmented or not, is that most algebras in such varieties are *not* graph algebras. For example, the product of graph algebras is typically not a graph algebra. Thus in order to find an equational base for the variety generated by a graph algebra it is not sufficient to consider graph algebras alone (augmented or not). Our strategy will be as follows:

Step 1: Give a finite generator G^* of the variety.

Step 2: Give a description of (possibly) all augmented graph algebras in the variety.

Step 3: Give a finite set of equations true in G^* (the hoped-for equational base).

Step 4: Use the equations of Step 3 to find a normal form for all groupoid terms.

Step 5: Determine all equations between normal forms not derivable from the equations of Step 3 and show that they fail in G^*.

Note that if G has two connected components G_0 and G_1, $G=G_0+G_1$, then an equation holds in G^* if and only if it holds both in G_0 and in G_1.

3.2. Theorem: *Let* $\mathbf{V}$ *be the variety* $\langle P_2+L_2\rangle$.

(a) $G^*\in\mathbf{V}$ *if and only if all connected components of* G *are complete or bipartite-complete.*

(b) *The equations* (0)—(9) *form an equational base for* $\mathbf{V}$.

(c) *By using the equations* (1)—(9), *every term in which* ∞ *does not occur can be transformed into one of the normal forms*

(i) x *(one variable),*

(ii) $x_1(x_1x_1)(x_2x_2)\dots(x_nx_n)$, $n\geqq 1$,

(iii) $x_1(y_1x_1)(y_2x_2)\dots(y_nx_n)$, $n\geqq 1$, $\{x_1,\dots,x_n\}\cap\{y_1,\dots,y_n\}=\emptyset$, *where* $u_1u_2\dots u_n = (\dots(u_1u_2)u_3\dots)u_n$ *(association from the left). In type* (iii) *the variables* $x_1, \dots, x_n$ *will be referred to as bottom variables and* $y_1, \dots, y_n$ *will be referred to as top variables.*

(d) *A regular equation* $\tau=\sigma$ *is derivable from* (0)—(9) *if and only if both sides have the same normal form, the same leftmost variable, and the same top and bottom variables (in the case of type* (iii)*) or both sides have an occurrence of* ∞.

(e) *An equation not derivable from* (0)—(9) *fails in* $(P_2+L_2)^*$.

Proof. The "if" direction of (a) is clear by previous reasoning and the "only if" direction follows from the rest of the theorem, since by Proposition 2.1 every other graph is inherently nonfinitely based and hence cannot be a member of the locally finite and finitely based variety **V**.

The condition (b) follows from (c)—(e) and the fact that (0)—(9) hold in $(P_2+L_2)^*$.

For (c): First we show that the terms of the form $x(u_1v_1)\dots(u_nv_n)$ $(n\geqq 0)$ are closed under multiplication (in the presence of (0)—(10)): Write $\varrho=x(u_1v_1)\dots(u_nv_n)$ and $\sigma=y(r_1s_1)\dots(r_ms_m)$. By using (3) repeatedly we have $\varrho\sigma=x[y(r_1s_1)\dots(r_ms_m)](u_1v_1)\dots(u_nv_n)$. If $m=0$, we use (4) to replace xy by $x(yx)$, and we are done. If $m>0$ then

$$x[y(r_1s_1)\dots(r_ms_m)] \underset{(6)}{=} x[y(r_1s_1)\dots(r_{m-1}s_{m-1})r_m](s_mr_m) \underset{(3)}{=}$$

$$\underset{(3)}{=} x[yr_m(r_1s_1)\dots(r_{m-1}s_{m-1})](s_mr_m) = \dots = x[yr_1\dots r_m](s_1r_1)\dots(s_mr_m).$$

By an induction on m using (5), (3) and (1), we obtain that $x[yr_1\dots r_m]=x(yr_1)\dots(yr_m)$ for $m\geqq 1$. Thus $\sigma\varrho$ is reduced to the desired form.

Since terms equivalent to terms of the desired form include the variables and are closed under multiplication, we conclude that any term τ without ∞ can be written in the form $x(u_1v_1)\dots(u_nv_n)$, where $x, u_1, v_1, \dots, u_n, v_n$ are the variables occurring in τ.

By equation (7) we can freely interchange the $v_1, \dots, v_n$ and then by (3) and (7) together we also can interchange the $u_1, \dots, u_n$ among each other. If some $u_i=v_j$ we can use (3) and (7) to obtain $u_1=v_1$ and then thence

$$x(u_1u_1)(u_2v_2)\dots(u_nv_n) = x(u_1u_1)(u_2u_2)(v_2v_2)\dots(u_nu_n)(v_nv_n),$$

as follows: $x(u_1u_1)(u_2u_2)(v_2v_2)\underset{(9)}{=}([x(u_1u_1)][x(u_1u_1)])(u_2v_2)$, and the computation of multiplicative closure given above shows that

$$[x(u_1u_1)][x(u_1u_1)] = x(xu_1)(u_1u_1) \underset{(8)}{=} x(u_1u_1)(u_1u_1) \underset{(10)}{=} x(u_1u_1).$$

Moreover, by (10), $x(u_1u_1)=x(xx)(u_1u_1)$, and therefore we may assume $x=u_1$. By (1) we may also assume that all the variables are distinct. If some $u_i=x$, we can use (8) to obtain $u_i=v_i$ for some i. Hence in these cases we arrive at normal form (ii).

Now we may assume $\{u_1, \ldots, u_n\}\cap\{v_1, \ldots, v_n\}=\emptyset$ and $x\notin\{u_1, \ldots, u_n\}$. By (2) and (4), $x(u_1v_1)=(xu_1)(u_1v_1)=x(u_1x)(u_1v_1)$, so we may assume $x\in\{v_1, \ldots, v_n\}$, say $x=v_1$. Thus in this case τ reduces to normal form (iii). It is not always possible to remove all duplications of variables; however, if there are duplications both among the top variables and among the bottom variables we can remove them using (3), (7) and (1). Furthermore, the following reasoning shows that duplications among the $u_1, \ldots, u_n$ can be arranged so that only u_1 is duplicated, and similarly for $v_1, \ldots, v_n$:

$$x(u_1v_1)(u_1v_2)(u_3v_3) \underset{(1)}{=} x(u_1v_1)(u_1v_2)(u_3v_3)(u_3v_3) \underset{(7)}{=}$$

$$\underset{(7)}{=} x(u_1v_3)(u_1v_3)(u_3v_1)(u_3v_2) \underset{(1)}{=} x(u_1v_3)(u_3v_1)(u_3v_2) \underset{(7)}{=} x(u_1v_1)(u_3v_2)(u_3v_3).$$

This reasoning already provides a proof for the "if" direction of (d), while the "only if" direction follows from (e).

For (e): By Lemma 3.1 we need consider only equations $\sigma=\infty$ and regular equations $\sigma=\tau$ in which ∞ does not occur.

Case 1: Let $\sigma=x$ be a regular equation with σ of type (ii) or (iii). Then $\sigma= =x(xx)=xx$ by (1) and (4) and hence the equation $\sigma=x$ is equivalent to the idempotent law $x=xx$, which holds in L_2^* but fails in P_i^* $(i=1, 2, 3)$.

Case 2: Let $\sigma=\tau$ be a regular equation with σ and τ of type (ii). If $\sigma=\tau$ is not derivable from (0)—(9) then σ and τ must have different leading variables x and y. Replacing all other variables by x we derive the equation $x(xx)(yy)= =y(yy)(xx)$, which is equivalent to $x(yy)=y(xx)$, by (10). The equation $x(yy)= =y(xx)$ clearly implies every regular equation $\sigma=\tau$ in which both σ and τ are of type (ii). $x(yy)=y(xx)$ is true in P_i^* $(i=1, 2, 3)$ and in L_1^* but not in L_2^*.

Case 3: Let $\sigma=\tau$ be a regular equation with σ of type (iii) and τ of type (ii). Substitute x for each bottom variable and y for each top variable of σ. Then $\sigma= =x(yx)=xy$ and $\tau=x(xx)(yy)=x(yy)$ or $\tau=y(xx)(yy)=y(xx)$ (see Case 2), and hence we can derive $xy=x(yy)$ or $xy=y(xx)$. From $xy=x(yy)$ we can derive the associative law:

$$x(yz) \underset{(2)}{=} (xy)(yz) = x(yy)(yz) \underset{(10)}{=} x(yy)(yy)(zz) = x(yy)(zz) = (xy)(zz) \underset{(3)}{=}$$

$$\underset{(3)}{=} (x(zz))y = (xz)y \underset{(3)}{=} (xy)z$$

and conversely the associative law implies $xy\underset{(1)}{=}(xy)y=x(yy)$. $xy=x(yy)$ holds in P_1^* and in L_2^* but not in P_2^* and L_3^*. On the other hand, from $xy=y(xx)$ we can derive the commutative law:

$$xy = y(xx) \underset{(1)}{=} (y(xx))(xx) \underset{(8)}{=} (y(yx))(xx) \underset{(6)}{=} y(y(xx)) = y(xy) \underset{(4)}{=} yx.$$

Then $xy=yx=x(yy)$, as before.

Case 4: Let $\sigma=\tau$ be a regular equation with σ, τ of type (iii). If σ and τ have different top and bottom variables the same substitution as in Case 3 will lead to an equation $xy=x(yy)\ldots$ or $xy=y(xx)\ldots$ and we are in Case 3. Since $\sigma=\tau$ is not derivable from (0)—(9) and since we assume that top and bottom variables coincide, the leading variables must be different, say x and z. Substitute y for all top variables and x for all bottom variables different from z and obtain from $\sigma=\tau$ the equation

$$x(yz) \underset{(2)}{=} (xy)(yz) = x(yx)(yz) = z(yx)(yz) \underset{(2)}{=} z(yx)$$

which in turn obviously implies $\sigma=\tau$. $x(yz)=z(yx)$ fails in P_3^* and L_2^* but holds in P_2^* and L_1^*.

Case 5: Let $\sigma=\infty$ be an equation such that ∞ does not occur in σ. By substituting x for all variables in σ we obtain the equation $xx=\infty$ or even $x=\infty$. These equations fail to hold in L_2^*. If σ has a normal form of type (iii), by substituting x for all the bottom variables and y for all the top variables, we can obtain $xy \underset{(4)}{=} x(yx) \underset{(2)}{=} \sigma=\infty$. Note that $x=\infty \Rightarrow xy=\infty \Rightarrow xx=\infty$, and moreover that $x=\infty \Leftrightarrow$ $\Leftrightarrow x=y$, $xy=\infty \Leftrightarrow xy=zz$, and $xx=\infty \Leftrightarrow xx=yy$.

This completes the proof of (e) and the whole theorem.

The preceding proof was a bit more elaborate than actually needed because we want next to classify the subvarieties of **V** and therefore need a classification of all possible equations, as given in the proof.

4. The lattice of finitely based subvarieties

4.1. Theorem. *The lattice of all subvarieties of* $\mathbf{V}=\langle P_2+L_2\rangle$ *is as given in Figure 6.*

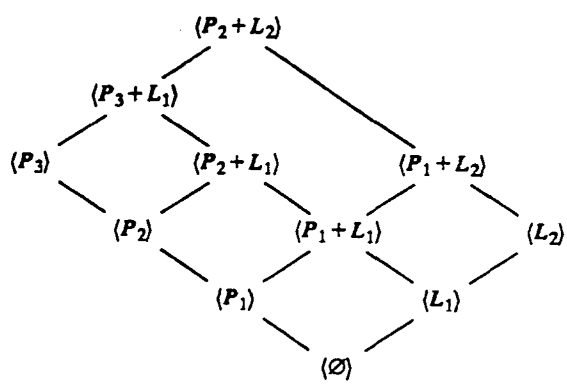


Figure 6

These varieties have the following equational bases:

$\langle P_2+L_2\rangle$: (0)—(9);

$\langle P_3+L_1\rangle$: (0)—(9), $x(yy)=y(xx)$;

$\langle P_3\rangle$: (0)—(9), $xx=yy$;

$\langle P_2\rangle$: (0)—(9), $xx=yy$, $x(yz)=z(yx)$;

$\langle P_1\rangle$: (0), $xy=uz$;

$\langle\emptyset\rangle$: $x=y$;

$\langle P_2+L_1\rangle$: (0)—(9), $x(yz)=z(yx)$;

$\langle P_1+L_2\rangle$: (0)—(9), $x(yz)=(xy)z$;

$\langle P_1+L_1\rangle$: (0), $x(yz)=(xy)z$, $xy=yx$, $xy=x(yy)$;

$\langle L_2\rangle$: (0), $x(yz)=(xy)z$, $xx=x$, $x(yz)=x(zy)$;

$\langle L_1\rangle$: (0), $x(yz)=(xy)z$, $xx=x$, $xy=yx$.

Proof. From the proof of Theorem 3.1 we can deduce the following classification of equations not derivable from (0)—(9).

(a) Each equation $\sigma=\infty$ not derivable from (0)—(9) is equivalent to one of $x=y$, $xy=zz$, or $xx=yy$.

(b) Each regular equation $x=\tau$ is equivalent to $xx=x$.

(c) Each regular equation $\sigma=\tau$ with both sides of type (ii) is equivalent to $x(yy)=y(xx)$.

(d) Each regular equation $\sigma=\tau$ with both sides of type (iii) with the same top and bottom variables is equivalent to $x(yz)=z(yx)$.

(e) Each of the remaining regular equations $\sigma=\tau$ implies $xy=x(yy)$ or even $xy=y(xx)$. The first of these implies every regular equation $\sigma=\tau$ with the same leading variable, because $xy=x(yy)\Leftrightarrow x(yz)=(xy)z$. The second of these implies every regular equation $\sigma=\tau$, because $xy=y(xx)\Leftrightarrow(x(yz)=(xy)z\ \&\ xy=yx)$.

This reasoning shows that every equation not derivable from (0)—(9) is equivalent in the presence of (0)—(9) to one of the equations $x=\infty$, $xy=\infty$, $xx=\infty$, $xx=x$, $x(yy)=y(xx)$, $xy=x(yy)$, $xy=y(xx)$, $x(yz)=z(yx)$. Moreover, between these we have the implications necessary to yield the diagram claimed by the theorem, when it is further observed that

$$xx=\infty \text{ and } xy=x(yy) \text{ imply } xy=\infty,$$
$$xx=\infty \text{ and } xx=x \text{ imply } x=\infty, \text{ and}$$
$$x(yy)=y(xx) \text{ and } xx=x \text{ imply } xy=yx.$$

4.2. Corollary. *For any graph G either G^* is inherently nonfinitely based or else G^* is finitely based and generates one of the eleven varieties in Theorem* 4.1.

5. Graph algebras versus augmented graph algebras

The situation for augmented graph algebras is now clear. To obtain similar results for graph algebras only a little more work is needed. By removing (0) from all the bases given in Theorem 4.1 and taking $[G]$ to mean the variety generated by $G^{\sharp}$, analogously to $\langle G\rangle$ for G^*, the arguments given above yield eleven finitely based subvarieties of $[P_2+L_2]$.

However, the analysis of the nonregular equations not derivable from (1)—(9) must now be done without the help of 3.1 and (0). It turns out that there are only three additional nonregular equations: $x=xy$, $xx=xy$, and $xx=x(yy)$. In the presence of (0) and in all graph algebras $x=xy$ is equivalent to $x=y$ and $xx=xy$ is equivalent to $xy=zz$. In the presence of (0) and (4) and in all graph algebras $xx=x(yy)$ is equivalent to $xx=yy$. However, (1)—(9) are not sufficient to establish any of these equivalences. There are, in fact, three additional subvarieties of $[P_2+L_2]$:

U_0, based on $x = xy$,

U_1, based on $xx = xy$,

U_2, based on (1)—(9) and $xx = x(yy)$.

These new varieties are not generated by graph algebras. In the case of graph algebras we obtain the lattice of Figure 7.

We deduce an analogue of Corollary 4.2 for graph algebras:

5.1. Corollary. *Let G be any graph. If G has an induced subgraph isomorphic to one of M, T, P_4, or L_3, then $[G]$ is inherently nonfinitely based. Otherwise, $[G]$ is one of the eleven finitely based varieties generated by graph algebras and appears in Figure 7.*

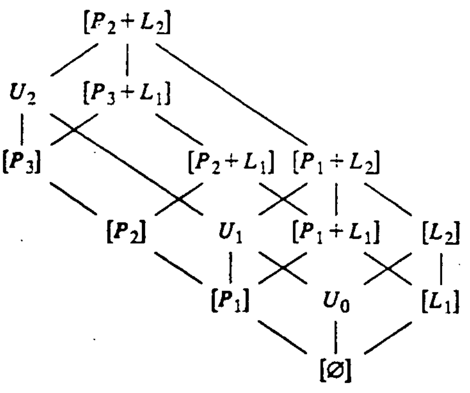


Figure 7

References

[1] K. A. BAKER, G. MCNULTY and H. WERNER, Shift-automorphism methods for inherently nonfinitely based varieties of algebras, preprint.

[2] S. BURRIS and H. P. SANKAPPANAVAR, *A Course in Universal Algebra,* Graduate Texts in Mathematics, v. 78, Springer-Verlag (New York, 1981).

[3] G. GRÄTZER, *Universal Algebra,* 2nd ed., Springer-Verlag (New York, 1979).

[4] J. JEŽEK, Nonfinitely based three-element idempotent groupoids, *Algebra Universalis,* **20** (1985), 292—301.

[5] B. JÓNSSON, *Topics in Universal Algebra,* Lecture Notes in Mathematics, vol. 250, Springer-Verlag (New York, 1972).

[6] E. KISS, A note on varieties of graph algebras, in: *Proceedings of the Charleston Conference on Universal Algebra and Lattice Theory* (S. Comer, ed.), Lecture Notes in Mathematics, vol. 1149, Springer-Verlag (New York, 1985); pp. 163—166.

[7] R. LYNDON, Identities in finite algebras, *Proc. Amer. Math. Soc.,* **5** (1954), 8—9.

[8] G. MCNULTY, How to construct finite algebras which are not finitely based, in: *Proceedings of the Charleston Conference on Universal Algebra and Lattice Theory* (S. Comer, ed.), Lecture Notes in Mathematics, vol. 1149, Springer-Verlag (New York, 1985); pp. 167—174.

[9] G. MCNULTY and C. SHALLON, Inherently nonfinitely based finite algebras, in: *Universal Algebra and Lattice Theory* (R. Freese and O. Garcia, eds.), Lecture Notes in Mathematics, vol. 1004, Springer-Verlag (New York, 1983); pp. 206—231.

[10] V. L. MURSKIĬ, The existence in three-valued logic of a closed class with finite basis not having a finite complete set of identities, *Dokl. Akad. Nauk. SSSR,* **163** (1965), 815—818; English Translation: *Soviet Math. Dokl.,* 6 (1965), 1020—1024.

[11] V. L. MURSKIĬ, On the number of k-element algebras with one binary operation without a finite basis of identities, *Problemy Kibernet.,* **35** (1979), 5—27.

[12] S. OATES-WILLIAMS, Graphs and universal algebras, in: *Combinatorial Mathematics VIII* (K. MacAvaney, ed.), Lecture Notes in Mathematics, vol. 884, Springer-Verlag (New York, 1981); pp. 351—354.

[13] P. PERKINS, Basis questions for general algebras, *Algebra Universalis,* **19** (1984), 16—23.

[14] R. PÖSCHEL, Graph varieties, preprint.

[15] R. PÖSCHEL, Graph algebras and graph varieties, preprint.

[16] R. PÖSCHEL and W. WESSEL, Classes of graphs which can be defined by equations in their graph algebras, preprint.

[17] C. SHALLON, *Nonfinitely Based Finite Algebras Derived from Lattices,* Ph. D. dissertation, UCLA, 1979.

(K. A. B.)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CA 90024, USA

(G. F. M.)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTH CAROLINA
COLUMBIA, SC 29208, USA

(H. W.)
FB 17 MATHEMATIK
GESAMTHOCHSCHULE
3500 KASSEL, BRD

Acta Sci. Math., 51 (1987), 17—33

Horn sentences in submodule lattices

GÁBOR CZÉDLI

To the memory of András P. Huhn

1. Introduction. Given a ring R with 1, a lattice is said to be representable by R-modules if it is embeddable in the lattice of submodules of some R-module. The class $\mathbf{L}(R)$ of all lattices representable by R-modules is known to be a quasivariety, i.e., to be axiomatizable by universal Horn sentences (cf. HERRMANN and POGUNTKE [9], HUTCHINSON [11] and, for another proof, [3]). The study of these quasivarieties was started in HUTCHINSON [10]. The main problem in this theory is to classify the possible quasivarieties of the form $\mathbf{L}(R)$. This needs to answer the following question:

(1.1) When does the inclusion $\mathbf{L}(R_1)\subseteq\mathbf{L}(R_2)$ hold?

Denoting by R-**Mod** $(\varkappa)$ the category of R-modules with cardinality less than or equal to a given cardinal $\varkappa$, the main result of [10] is the following.

Theorem 1.2 (HUTCHINSON [10]). $\mathbf{L}(R_1)\subseteq\mathbf{L}(R_2)$ *if and only if for each infinite cardinal* $\varkappa$ *there exists an exact embedding functor* R_1-**Mod** $(\varkappa)\to R_2$-**Mod**.

Note that even a stronger result (cf. HUTCHINSON [11B]) is true: $\mathbf{L}(R_1)\subseteq\mathbf{L}(R_2)$ iff there is an exact embedding functor R_1-**Mod**$\to R_2$-**Mod**.

By the help of this theorem, HUTCHINSON [10] proves a number of interesting results concerning (1.1). As the proof and the applications of this theorem require a good command of category theory and a hard technique, it seems reasonable to develop another approach to (1.1). As $\mathbf{L}(R)$ is a quasivariety, the inclusion $\mathbf{L}(R_1)\subseteq\subseteq\mathbf{L}(R_2)$ holds if and only if every Horn sentence satisfied in $\mathbf{L}(R_2)$ is also satisfied in $\mathbf{L}(R_1)$. Therefore (1.1) can be reduced to the following problem:

(1.3) When does a Horn sentence hold in $\mathbf{L}(R)$?

Received February 7, 1986.

Research partially supported by Hungarian National Foundation for Scientific Research grant no. 1813.

Our aim in the present paper is to investigate the connection between properties of rings R and Horn sentences holding in $\mathbf{L}(R)$. We give some answer to (1.3) in Theorem 3.5, which, among others, enables us to give new proofs for some results of HUTCHINSON [10] concerning (1.1). Although our description (Theorems 4.1 and 4.2) of the ring properties that can be characterized by Horn sentences is not complete, it leads to a solution of the following problem of JÓNSSON [13]:

(1.4) Is there a strong Mal'tsev condition for any Horn sentecne χ which characterizes if χ holds in the congruence lattices of algebras of an n-permutable variety?

The connection between ring properties and lattice identities, which are particular Horn sentences, was firstly studied by HERRMANN and HUHN [8]. After András P. Huhn had personally initiated me into their research with C. Herrmann, we with G. Hutchinson settled the case of lattice identities in [12]. The present paper resembles [12] in some extent; e.g., the use of Mal'tsev conditions is the main tool of investigations in both papers. The results of this paper are taken from the author's thesis [4].

2. Preliminaries. By a ring we always mean a ring with 1, and modules are always unitary left modules. Suppose R is a ring, let R-**Mod** denote the class of R-modules. If M is an R-module then Con (M) and Su (M) will stand for the lattice of congruences and that of submodules of R, respectively. For a class $\mathcal{M}$ of modules, let $\mathrm{Con}\,(\mathcal{M})=\{\mathrm{Con}\,(M)\colon M\in\mathcal{M}\}$ and $\mathrm{Su}\,(\mathcal{M})=\{\mathrm{Su}\,(M)\colon M\in\mathcal{M}\}$. Then $\mathbf{L}(R)=\mathbf{IS}\,\mathrm{Su}\,(R\text{-}\mathbf{Mod})$. As $\mathrm{Con}\,(M)\cong\mathrm{Su}\,(M)$ for any $M\in R$-**Mod** (cf. BIRKHOFF [1, p. 159]), we have $\mathbf{L}(R)=\mathbf{IS}\,\mathrm{Con}\,(R\text{-}\mathbf{Mod})$. It is worth pointing out that exactly the same Horn sentences hold in $\mathbf{L}(R)$, Su (R-**Mod**) and Con (R-**Mod**), whence, in many of the forthcoming results, $\mathbf{L}(R)$ can be replaced by any of the other two. The lattice variety generated by $\mathbf{L}(R)$ will be denoted by $\mathbf{HL}(R)$, which consists of all homomorphic images of lattices in $\mathbf{L}(R)$.

For any integers m and n, let $D(m, n)$ denote the sentence (in the first-order language of rings with 1) "$(\exists x)(m\cdot x=n\cdot 1)$" where $k\cdot y$ or ky is an abbreviation for $y+y+\ldots+y$ (k times if $k>0$) or 0 (if $k=0$) or $-|k|\cdot y$ (if $k<0$). $D(m, n)$ is called a *divisibility condition.* Denoting the set of prime numbers by P, a map $S\colon \{0\}\cup P\to\omega+1$ is called a *spectrum* if

(α) $S(0)<\omega$

and

(β) if $S(0)>0$ then $S(p)=\max\{i\colon 0\leqq i$ and p^i divides $S(0)\}$ holds for all $p\in P$.

For any spectra S_1 and S_2, let $S_1\leqq S_2$ mean that $S_1(0)$ divides $S_2(0)$ and, for all $p\in P$, $S_1(p)\leqq S_2(p)$. Equipped with this (ordering) relation, the set $\mathcal{L}_S$ of all spectra turns into a complete lattice (cf. Theorem 2.1 later). For a ring R, let S_R be the map $\{0\}\cup P\to\omega+1$ defined by $S_R(0)=\operatorname{char} R=\min\{i\colon i\geqq 1$ and $D(0,i)$ holds in $R\}$, the characteristic of R (here $\min\emptyset=0$) and, for $p\in P$, $S_R(p)=\min\{i\colon 0\leqq i<\omega$ and $D(p^{i+1},p^i)$ holds in $R\}$ (here $\min\emptyset=\omega$). HUTCHINSON [12] has shown that S_R is a spectrum; it will be called the spectrum of R.

Now, for a spectrum S with $S(0)=0$, let $FR(\{x_p\colon p\in P\})$ be the free commutative ring with 1 on the free generating set $\{x_p\colon p\in P\}$, let J_S denote the ideal of this ring generated by $\{p^{S(p)}(px_p-1)\colon p\in P$ and $S(p)<\omega\}$, and put $R_S=FR(\{x_p\colon p\in P\})/J_S$. For $S(0)=m>0$, we put $R_S=\mathbf{Z}_m$, the factor ring of the ring $\mathbf{Z}$ of integers modulo m.

For an integer n and a prime p, let $\exp(n,p)$ denote $\sup\{i\colon 0\leqq i<\omega$ and p^i divides $n\}$. Then the main result of [12] is the following

Theorem 2.1 (HUTCHINSON [12]). (a) $\mathbf{HL}(R)$ *and* S_R *mutually determine each other.*

(b) *The lattice varieties of the form* $\mathbf{HL}(R)$, R *is a ring, form a complete lattice* $\mathcal{L}_R$ *under the inclusion.*

(c) $\mathcal{L}_R$ *is isomorphic to* $\mathcal{L}_S$. *In fact, the map* $\mathcal{L}_R\to\mathcal{L}_S$, $\mathbf{HL}(R)\mapsto S_R$ *is a lattice isomorphism whose inverse is* $\mathcal{L}_S\to\mathcal{L}_R$, $S\mapsto\mathbf{HL}(R_S)$.

(d) $D(0,n)$ *holds in a ring* R *iff* $S_R(0)$ *divides* n *while, for* $m\neq 0$, $D(m,n)$ *holds in* R *iff* $(\forall p\in P)(\exp(m,p)>\exp(n,p)\Rightarrow\exp(n,p)\geqq S_R(p))$.

By a Horn sentence we mean a universally quantified first order lattice sentence χ of the form

$$(2.2)\quad (p_0\leqq q_0\ \&\ p_1\leqq q_1\ \&\ \dots\ \&\ p_t\leqq q_t)\Rightarrow p\leqq q$$

where $-1\leqq t<\omega$ and $p_0,q_0,p_1,q_1,\dots,p_t,q_t,p,q$ are lattice terms. (In case $t=-1$ the premise is empty and χ is the identity $p\leqq q$.) Let us call χ *regular* if, for any two rings R_1 and R_2, $S_{R_1}=S_{R_2}$ and $\mathbf{L}(R_1)\models\chi$ imply $\mathbf{L}(R_2)\models\chi$. I.e., χ is regular iff the satisfaction of χ in $\mathbf{L}(R)$ depends only on S_R or, equivalently, on $\mathbf{HL}(R)$. By Theorem 2.1 (a), every lattice identity is regular. In Sections 4 and 8, we will deal with ring properties characterizable by regular Horn sentences as we have not succeeded in handling the general case. (This situation resembles [2].) HUTCHINSON [10] has shown that there are rings R_1 and R_2 such that $S_{R_1}=S_{R_2}$ but $\mathbf{L}(R_1)\neq\mathbf{L}(R_2)$, whence there exist irregular Horn sentences, too. In the forthcoming [6] we will explicitly construct an irregular Horn sentence.

3. Mal'tsev type conditions. Given an integer $n\geqq 2$ and a Horn sentence χ, [3] associates a Mal'tsev condition with χ such that the satisfaction of χ in the congruence lattices of an arbitrary n-permutable variety $\mathcal{V}$ is equivalent to the satisfaction of this Mal'tsev condition in $\mathcal{V}$. Unfortunately, the Mal'tsev conditions in [3] are so complicated that instead of recalling them and adapting them to the special case $\mathscr{U}=R$-**Mod** it is better and shorter to develop them independently. As these conditions will be meaningful only when $\mathscr{U}=R$-**Mod**, they will be referred to as Mal'tsev *type* conditions.

Our Mal'tsev type conditions will be given by certain graphs. First, for any lattice term $p=p(x\colon x\in U)$ we define a graph $G(p)$ associated with p. (Here we adopt the abbreviation $p(x\colon x\in\{x_1, x_2, \ldots, x_n\})$ for $p(x_1, \ldots, x_n)$. U is assumed to have a fixed order.) The edges of $G(p)$ will be coloured by the variables $x\in U$, and two distinguished vertices, the so-called left and right endpoints, will have special roles. In figures these endpoints will always be placed on the left-hand side and on the right-hand side, respectively. An x-coloured edge connecting the vertices u and v will often be denoted by (u, x, v). Before defining $G(p)$ we introduce two kinds of operations for graphs. We obtain the *parallel connection* of graphs G_1 and G_2 by taking disjoint copies of G_1 and G_2 and identifying their left (right, resp.) endpoints (Figure 3.1).

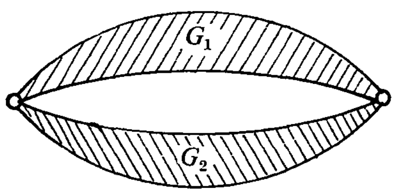


Figure 3.1

Figure 3.2

Similarly, we obtain the *serial connection* of G_1 and G_2 by taking disjoint copies of G_1 and G_2 and identifying the right endpoint of G_1 and the left endpoint of G_2. (The left endpoint of G_1 and the right endpoint of G_2 are the endpoints of the serial connection, cf. Figure 3.2.) Now if p is a variable then $G(p)$ is the following graph

$$\circ\!\!\xrightarrow{\;\;p\;\;}\!\!\circ\,,$$

which consists of a single edge coloured by p. Let $G(p_1\wedge p_2)$ $(G(p_1\vee p_2)$, resp.) be the parallel connection (serial connection, resp.) of the graphs $G(p_1)$ and $G(p_2)$. This defines $G(p)$ for any lattice term p via induction on the length of p. For a graph G, let $V(G)$ and $E(G)$ denote the vertex set and the edge set of G, respectively. Note that $E(G(p))\subseteq V(G(p))\times U\times V(G(p))$ if $p=p(x\colon x\in U)$.

Now let $p=p(x\colon x\in U)$ be a lattice term, let R be a ring, let $M\in R$-**Mod**, and let φ be a map from U into $\mathrm{Su}(M)$. A map $\psi\colon V(G(p))\to M$ will be called a *connecting map* (with respect to φ) if (left endpoint)$\psi=0$ and $b\psi-a\psi\in x\varphi$ holds for every edge $(a, x, b)\in E(G(p))$. For a graph G, let $(\varphi, \psi)\colon G\xrightarrow{c} M$ denote the fact that $\psi\colon V(G)\to M$ is a connecting map with respect to φ. Given a $y\in M$, if there exists a connecting map $\psi\colon V(G(p))\to M$ such that (right endpoint)$\psi=y$ then y will be said to be *attainable* by $G(p)$ (with respect to φ). Knowing that $X\vee Y= =X+Y=\{x+y\colon x\in X \text{ and } y\in Y\}$ and $X\wedge Y=X\cap Y$ hold for $X, Y\in \mathrm{Su}(M)$, an easy (and therefore omitted) induction on the length of p yields the following

Lemma 3.3. *For any $y\in M$, $y\in p(x\varphi\colon x\in U)$ iff y is attainable by $G(p)$ with respect to φ.*

The following lemma will also be useful.

Lemma 3.4. *Assume that $t(x\colon x\in U)$ is a lattice term, M and K are modules over a ring R, $\psi\colon M\to K$ is a homomorphism, $\mu\colon U\to \mathrm{Su}(M)$ and $\varphi\colon U\to \mathrm{Su}(K)$ are maps, and $x\mu\psi\subseteq x\varphi$ for all $x\in U$. Then $t(x\mu\colon x\in U)\psi\subseteq t(x\varphi\colon x\in U)$.*

Proof. The proof goes via induction on the length of t. If $t\in U$, i.e. t is a variable, then the statement is obvious. If the statement is already true for t_1 and t_2 then for $t=t_1\vee t_2$ we have

$$t(x\mu\colon x\in U)\psi = (t_1(x\mu\colon x\in U)+t_2(x\mu\colon x\in U))\psi =$$
$$= t_1(x\mu\colon x\in U)\psi+t_2(x\mu\colon x\in U)\psi \subseteq t_1(x\varphi\colon x\in U)+t_2(x\varphi\colon x\in U) = t(x\varphi\colon x\in U),$$

while in the case $t=t_1\wedge t_2$ we have

$$t(x\mu\colon x\in U)\psi = (t_1(x\mu\colon x\in U)\cap t_2(x\mu\colon x\in U))\psi \subseteq$$
$$\subseteq t_1(x\mu\colon x\in U)\psi\cap t_2(x\mu\colon x\in U)\psi \subseteq t_1(x\varphi\colon x\in U)\cap t_2(x\varphi\colon x\in U) = t(x\varphi\colon x\in U).$$

If G is a graph and H is a set then let $H\times G$ denote the graph whose vertex set and edge set are $H\times V(G)$ and $\{((h,a),x,(h,b))\colon (a,x,b)\in E(G)\}$, respectively. Note that $H\times G$ is isomorphic to $\dot{\bigcup}_{h\in H} G$, the disjoint union of $|H|$ copies of G.

Let us fix a ring R and a Horn sentence χ of the form (2.2) where $t\geqq 0$. (The assumption $t\geqq 0$ does not hurt the generality as any lattice identity $p\leqq q$ is equivalent, modulo lattice theory, to the Horn sentence $x\leqq x\Rightarrow p\leqq q$.) Let U be the set of variables occurring in χ. Before formulating Theorem 3.5, we have to define certain modules over R. It seems reasonable to outline our goal roughly before the following tedious definition. In order to obtain a necessary condition for the satisfaction of χ in $\mathbf{L}(R)$ we will start from a "small" module M^0, submodules X^0 for $x\in U$, and an element $f_1\in p(X^0\colon x\in U)$. If $p_0(X^0\colon x\in U)\leqq q_0(X^0\colon x\in U)$ fails then, in order to improve this failure, we will extend X^0, $x\in U$, and M^0 to appropriate X^1 and M^1, respectively. Then, by extending X^1, $x\in U$, and M^1 to X^2 and M^2 if necessary, we will try to remedy the failure of $p_1(X^1\colon x\in U)\leqq q_1(X^1\colon x\in U)$; etc. After ω steps we will obtain $M^\omega=\dot{\bigcup}_{m<\omega} M^m$ and, for $x\in U$, $X^\omega=\bigcup_{m<\omega} X^m$. Now the premise of χ will hold for X^ω, $x\in U$, and the satisfaction of χ in $\mathbf{L}(R)$ will imply $f_1\in q(X^\omega\colon x\in U)=\bigcup_{m<\omega} q(X^m\colon x\in U)$. Lemma 3.3 will be our main tool in doing so.

Now the precise definition comes. First we define lattice terms p_i and q_i for $t<i<\omega$: let p_i and q_i be p_j and q_j, respectively, where $j\equiv i \mod (t+1)$ and $0\leqq j\leqq t$. For any integer $m\geqq 0$, we intend to define a graph G^m, a subset F^m of $V(G^m)$, an R-module M^m and submodules X^m of M^m (for all $x\in U$) by induction such that $V(G^m)\subseteqq M^m$, M^m is freely generated by F^m and, for all $x\in U$, X^m is the submodule of M^m generated by $\{c-b\colon (b, x, c)\in E(G^m)\}$, in notation $X^m=[c-b\colon (b, x, c)\in E(G^m)]$. (Here we have a map $U\to \mathrm{Su}(M^m)$ which we denote by capitalizing and adding a superscript, e.g. $x\mapsto X^m$ and $y_i\mapsto Y_i^m$ for $x, y_i\in U$.) As G^m and F^m will determine M^m and X^m, $x\in U$, it will suffice to define the former two.

Let $G^0:=G(p)$ and $F^0:=V(G(p))\setminus\{\text{left endpoint}\}$, and, in order to ensure $V(G^0)\subseteqq M^0$, identify the left endpoint of G^0 and the zero of M^0.

Assume that G^{m-1}, F^{m-1}, M^{m-1} and X^{m-1}, $x\in U$, have already been defined for some $m\geqq 1$. Now the definition ramifies as we want to define two kinds of our graphs and modules.

(a) Choose a subset S_m of M^{m-1} such that $S_m\subseteqq P_m^{m-1}$ where $P_m^{m-1}= =p_m(X^{m-1}\colon x\in U)$.

(b) Choose a subset S_m of M^{m-1} such that $P_m^{m-1}\setminus Q_m^{m-1}\subseteqq [S_m]\subseteqq P_m^{m-1}$ where P_m^{m-1}, Q_m^{m-1} and $[S_m]$ denote $p_m(X^{m-1}\colon x\in U)$, $q_m(X^{m-1}\colon x\in U)$ and the submodule generated by S_m.

In both cases, we put

$$F^m:=F^{m-1}\cup(\{m\}\times S_m\times(V(G(q_m))\setminus\{\text{left endpoint, right endpoint}\})).$$

We obtain G^m from $G^{m-1}\cup(\{m\}\times S_m\times G(q_m))$ by identifying the zero of M^m and all the $(m, s, \text{left endpoint})$, $s\in S_m$, and by identifying $(m, s, \text{right endpoint})$ and s for every $s\in S_m$. Then $V(G^{m-1})\subseteqq V(G^m)\subseteqq M^m$ and G^{m-1} is a (weak) subgraph of G^m, i.e., $E(G^{m-1})\subseteqq E(G^m)\cap(V(G^{m-1})\times U\times V(G^{m-1}))$. Therefore $X^{m-1}\subseteqq X^m$, $x\in U$. Obviously, $F^{m-1}\subseteqq F^m$ and $M^{m-1}\subseteqq M^m$.

Now we have defined G^m, F^m, M^m and X^m, $x\in U$, for all $m\geqq 0$. Note that, in both cases, these things depend on the choice of $S_1, S_2, S_3, \dots$ because we want to make the following theorem easy to handle. We also note that the choice $S_1=P_1^0$, $S_2=P_2^1$, $S_3=P_3^2$, ... is always possible. Let f_1 denote the right endpoint of $G^0=G(p)$, then we have

Theorem 3.5. (A) *Suppose that* $S_1, S_2, S_3, \dots$ *are chosen according to* (a). *If there exists a non-negative integer* n *such that* $f_1\in q(X^n\colon x\in U)$ *then* χ *holds in* $\mathbf{L}(R)$ *or, equivalently, in* $\mathrm{Su}(R\text{-}\mathbf{Mod})$.

(B) *Suppose that* $S_1, S_2, S_3, \dots$ *are choosen according to* (b). *Then* χ *holds in* $\mathbf{L}(R)$ *if and only if there exists a non-negative integer* n *such that* $f_1\in q(X^n\colon x\in U)$.

Proof. It suffices to prove (A) and the "only if" part of (B).

To prove (A), assume that $f_1\in q(X^n\colon x\in U)$ holds for some n. Let $A\in R\text{-}\mathbf{Mod}$, for $x\in U$ let $X'\in\mathrm{Su}(A)$, let $a_1\in p(X'\colon x\in U)$, and assume that $p_i(X'\colon x\in U)\subseteqq$

$\subseteqq q_i(X'\colon x\in U)$ holds for $i\leqq t$ (whence for $i<\omega$ as well). Let φ denote the map $U\to \mathrm{Su}\,(A)$, $x\mapsto X'$. We need to show $a_1\in q(X'\colon x\in U)$. Via induction on m, we intend to define two maps, $\eta^m\colon V(G^m)\to A$ and $\psi^m\colon M^m\to A$ for any $m\geqq 0$ such that

(I_m) $(\varphi, \eta^m)\colon G^m \xrightarrow{c} A$, and $\psi^m\colon M^m\to A$ is a homomorphism extending both η^m and η^0.

By Lemma 3.3, a_1 is attainable by $G(p)=G^0$ with respect to φ. I.e., there is a map $\eta^0\colon V(G^0)\to A$ such that $f_1\eta^0=a_1$ and $(\varphi, \eta^0)\colon G^0\xrightarrow{c} A$. Extend $\eta^0{\restriction}F^0$ to a homomorphism $\psi^0\colon M^0\to A$. (Here $\restriction$ stands for the restriction.) As M^0 is freely generated by F^0, ψ^0 exists and is uniquely determined. Since $0\eta^0=(\text{left endpoint})\eta^0=0$, ψ^0 extends η^0, too, and (I_0) is satisfied.

Now let $m\geqq 1$ and suppose (I_{m-1}). I.e., $(\varphi, \eta^{m-1})\colon G^{m-1}\xrightarrow{c} A$ and ψ^{m-1} extends both η^{m-1} and η^0. For $x\in U$,

$$X^{m-1}\psi^{m-1}=[c-b\colon (b,x,c)\in E(G^{m-1})]\psi^{m-1}=$$
$$=[c\psi^{m-1}-b\psi^{m-1}\colon (b,x,c)\in E(G^{m-1})]=[c\eta^{m-1}-b\eta^{m-1}\colon (b,x,c)\in E(G^{m-1})]\subseteqq X',$$

whence, by Lemma 3.4,

$$S_m\psi^{m-1}\subseteqq P_m^{m-1}\psi^{m-1}=p_m(X^{m-1}\colon x\in U)\psi^{m-1}\subseteqq p_m(X'\colon x\in U)\subseteqq q_m(X'\colon x\in U).$$

I.e., $S_m\psi^{m-1}\subseteqq q_m(X'\colon x\in U)$. By Lemma 3.3, for every $s\in S_m$, $s_m\psi^{m-1}$ is attainable by $G(q_m)\cong\{m\}\times\{s\}\times G(q_m)$ with respect to φ. I.e., there is a map η_s^m such that $(\varphi,\eta_s^m)\colon \{m\}\times\{s\}\times G(q_m)\xrightarrow{c} A$ and $(m, s, \text{right endpoint})\eta_s^m=s\psi^{m-1}$. Put $\eta^m=$ $=\eta^{m-1}\cup\overline{\bigcup_{s\in S_m}}\eta_s^m$. Then η^m is really a map from $V(G^m)$ into A, and it extends η^{m-1}. Further, if $s\in S_m$ then $s\eta^m=(m, s, \text{right endpoint})\eta^m=s\psi^{m-1}$. Now let $\psi^m\colon M^m\to A$ be the unique homomorphism that extends $\eta^m{\restriction}F^m$. For any $u\in F^{m-1}\subseteqq F^m$, $u\psi^m=$ $=u\eta^m=u\eta^{m-1}=u\psi^{m-1}$. Hence $\psi^m{\restriction}F^{m-1}=\psi^{m-1}{\restriction}F^{m-1}$ and $[F^{m-1}]=M^{m-1}$ yield that ψ^m extends ψ^{m-1}. For $u\in V(G^m)\setminus F^m$ either $u\in V(G^{m-1})$ and $u\eta^m=u\eta^{m-1}=$ $=u\psi^{m-1}=u\psi^m$ or $u\in S_m$ and $u\eta^m=(m, u, \text{right endpoint})\eta^m=u\psi^{m-1}=u\psi^m$. Hence ψ^m is an extension of η^m. As η^{m-1} and η_s^m, $s\in S_m$, are connecting maps, so is η^m. I.e., $(\varphi,\eta^m)\colon G^m\xrightarrow{c} A$, and (I_m) holds.

Now η^m and ψ^m, satisfying (I_m), are defined for all $m\geqq 0$. From (I_n) we conclude that, for $x\in U$,

$$X^n\psi^n=[c-b\colon (b,x,c)\in E(G^n)]\psi^n=[c\psi^n-b\psi^n\colon (b,x,c)\in E(G^n)]=$$
$$=[c\eta^n-b\eta^n\colon (b,x,c)\in E(G^n)]\subseteqq X'.$$

Hence Lemma 3.4 yields $a_1=f_1\eta^0=f_1\psi^n\in q(X^n\colon x\in U)\psi^n\subseteqq q(X'\colon x\in U)$. This proves (A).

To prove the "only if" part of (B), assume that χ holds in $\mathbf{L}(R)$ and F^m, G^m, M^m, X^m $(0\leqq m<\omega,\ x\in U)$ are defined according to (b). Then $F^0\subseteqq F^1\subseteqq F^2\subseteqq\ldots$,

$M^0 \subseteq M^1 \subseteq M^2 \subseteq \ldots$ and $G^0 \subseteq G^1 \subseteq G^2 \subseteq \ldots$. Put $F^\omega := \bigcup_{m<\omega} F^m$ and $G^\omega := \bigcup_{m<\omega} G^m$ (i.e., $V(G^\omega) = \bigcup_{m<\omega} V(G^m)$ and $E(G^\omega) = \bigcup_{m<\omega} E(G^m)$). Let M^ω be the R-module freely generated by F^ω and, for $x \in U$, let $X^\omega = [c-b\colon (b, x, c) \in G^\omega]$. It is easy to see that $V(G^\omega) \subseteq M^\omega$, $M^\omega = \bigcup_{m<\omega} M^m$ and, for $x \in U$, $X^\omega = \bigcup_{m<\omega} X^m$. We will show that the submodules X^ω, $x \in U$, satisfy the premise of χ.

Let the map $U \to \mathrm{Su}\,(M^l)$, $x \mapsto X^l$ be denoted by φ^l, $l \leqq \omega$. Since $X^\omega = \bigcup_{m<\omega} X^m$ and $X^0 \subseteq X^1 \subseteq X^2 \subseteq \ldots$, we obtain that, for any map $\eta\colon V(G(p_j)) \to M^\omega$, $(\varphi^\omega, \eta)\colon G(p_j) \xrightarrow{c} M$ iff there is an m such that $(\varphi^{m-1}, \eta)\colon G(p_j) \xrightarrow{c} M^\omega, j+1 < m < \omega$ and $j \equiv m \mod (t+1)$. Hence, denoting the right endpoint of $G(p_j)$ by r and applying Lemma 3.3, we obtain

$$
\begin{aligned}
(3.6) \qquad p_j(X^\omega\colon x \in U) &= \{b\colon (\exists\eta)(r\eta = b \text{ and } (\varphi^\omega, \eta)\colon G(p_j) \xrightarrow{c} M^\omega)\} = \\
&= \{b\colon (\exists\eta)(\exists m)(j+1 < m < \omega,\ j \equiv m \bmod (t+1),\ r\eta = b \\
&\qquad \text{and } (\varphi^{m-1}, \eta)\colon G(p_j) \xrightarrow{c} M^\omega)\} = \\
&= \bigcup(\{b\colon (\exists\eta)(r\eta = b \text{ and } (\varphi^{m-1}, \eta)\colon G(p_j) \xrightarrow{c} M^\omega)\}\colon \\
&\qquad j+1 < m < \omega \text{ and } j \equiv m \bmod (t+1)) = \\
&= \bigcup(p_j(X^{m-1}\colon x \in U)\colon j+1 < m < \omega \text{ and } j \equiv m \bmod (t+1)).
\end{aligned}
$$

Since $(\varphi^m, \text{identical map})\colon \{m\} \times \{s\} \times G(q_m) \xrightarrow{c} M^\omega$, Lemma 3.3 yields $s \in q_m(X^m\colon x \in U)$ for any $s \in S_m$, $m < \omega$. Therefore $[S_m] \subseteq q_m(X^m\colon x \in U)$. For $j+1 < m < \omega$ and $j \equiv m \mod (t+1)$ we obtain

$$
\begin{aligned}
p_j(X^{m-1}\colon x \in U) &= p_m(X^{m-1}\colon x \in U) = P_m^{m-1} = \\
&= (P_m^{m-1} \setminus Q_m^{m-1}) \cup (P_m^{m-1} \cap Q_m^{m-1}) \subseteq [S_m] \cup Q_m^{m-1} \subseteq \\
&\subseteq q_m(X^m\colon x \in U) \cup q_m(X^{m-1}\colon x \in U) \subseteq \\
&\subseteq q_m(X^\omega\colon x \in U) \cup q_m(X^\omega\colon x \in U) = q_m(X^\omega\colon x \in U) = q_j(X^\omega\colon x \in U).
\end{aligned}
$$

This inclusion and (3.6) yield $p_j(X^\omega\colon x \in U) \subseteq q_j(X^\omega\colon x \in U)$, whence the premise of χ holds for X^ω, $x \in U$. As $\mathrm{Su}\,(M^\omega) \in \mathbf{L}(R)$, $p(X^\omega\colon x \in U) \subseteq q(X^\omega\colon x \in U)$. Lemma 3.3 yields $f_1 \in p(X^\omega\colon x \in U)$ as $(\varphi^\omega, \text{identical map})\colon G(p) \xrightarrow{c} M$. An argument analogous to (3.6) shows that $q(X^\omega\colon x \in U) = \bigcup_{m<\omega} q(X^m\colon x \in U)$. Hence we have $f_1 \in p(X^\omega\colon x \in U) \subseteq q(X^\omega\colon x \in U) = \bigcup_{m<\omega} q(X^m\colon x \in U)$. Therefore there is an n such that $f_1 \in q(X^n\colon x \in U)$, which completes the proof.

4. Regular Horn sentences. Let U denote the set $\{x, y, z, t, e\}$ of variables, and define the following lattice terms over U (the meet and join will be denoted by $\cdot$ and

+, respectively):

$$p := (x+y)(z+t), \quad w_{-1} := (x+z)(y+t), \quad w_0 := x,$$
$$s_{i+1} := (w_i+t)(y+z) \quad \text{and} \quad w_{i+1} := (s_{i+1}+p)(x+z) \quad \text{for} \quad i \geqq 0.$$

By induction, this defines s_i and w_j for all $i \geqq 1$, $j \geqq -1$. Now let m, n and k be non-negative integers, put

$$p_0 := \big((e+w_{n-1})w_{-1}+x\big)z, \quad q_0 := e,$$
$$q := \big((w_k+y)(z+t)+w_{mn}\big)(x+y+e)+x+z,$$

and let $\chi(m, n, k)$ denote the Horn sentence

$$p_0 \leqq q_0 \Rightarrow p \leqq q.$$

Theorem 4.1. *For any ring R and non-negative integers m, n, k, the Horn sentence $\chi(m, n, k)$ holds in $\mathbf{L}(R)$ if and only if there exists a non-negative integer i such that the divisibility condition $D(mn^{i+1}, kn^i)$ holds in R.*

Note that, in virtue of Theorems 2.1 (d) and 4.1, $\chi(m, n, k)$ is regular. To avoid the feeling that $(\exists i)\big(D(mn^{i+1}, kn^i)\big)$ in the above theorem is just a haphazard ring property we state the following result, which is almost the converse of Theorem 4.1. While we have collected all we need to prove Theorem 4.1, the following theorem will be proved only in Section 8.

Theorem 4.2. *Let χ be a regular Horn sentence. Assume that there is a ring R^* of characteristic 0, i.e., $S_{R^*}(0)=0$, such that χ holds in $\mathbf{L}(R^*)$. Then there are positive integers m_χ, n_χ and k_χ such that, for any ring R, χ holds in $\mathbf{L}(R)$ if and only if $D(m_\chi n_\chi^{i+1}, k_\chi n_\chi^i)$ holds in R for some integer $i \geqq 0$.*

Proof of Theorem 4.1. We will apply Theorem 3.5 (B) with the choice $S_j = p_j(H^{j-1}\colon h \in U)$. The graph $G^0 = G(p)$ is given in Figure 4.3, whence $X = X^0 = [f_2]$, $Y = Y^0 = [f_1 - f_2]$, $Z = Z^0 = [f_3]$ and $T = T^0 = [f_1 - f_3]$. Since $G(q_0) = G(q_j)$, $0 \leqq j < \omega$, has no "inner vertex", i.e., $|V(G(q_j))| = 2$, we have $F^j = F^0 = \{f_1, f_2, f_3\}$ and $M^j = M^0$, $0 \leqq j < \omega$. Let F and M denote F^0 and M^0, respectively. As the only edge of $G(q_j) = G(q_0)$ is coloured by e, all the edges in $E(G^j) \setminus E(G^0)$ are coloured

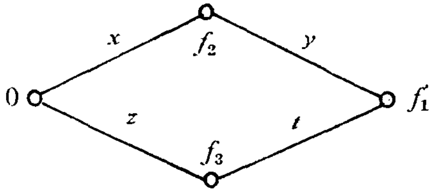


Figure 4.3

by e, and we have $X^j=X^0$, $Y^j=Y^0$, $Z^j=Z^0$, $T^j=T^0$, $E^0=\{0\}$ and $E^j=[E^{j-1}\cup p_0(H^{j-1}\colon h\in U)]=E^{j-1}+p_0(H^{j-1}\colon h\in U)$. We claim that

$$P=[f_1],\quad W_j=[f_2+jf_3]\quad\text{for}\quad j\geqq -1,$$
$$S_i=[f_1-f_2-if_3]\quad\text{for}\quad i\geqq 1,\quad\text{and}\tag{4.4}$$
$$E^j=P_0^{j-1}=\{rf_3\colon r\in R\text{ and } n^jr=0\}$$

where V stands for $v(H^0\colon h\in U)$ if $v\in\{p,w_j,s_i\}$. These formulas can be obtained by an elementary calculation, only a part of which will be presented. As any element a of M can uniquely be written of the form $a=r_1f_1+r_2f_2+r_3f_3$ where $r_1,r_2,r_3\in R$, we can compute as follows.

$$P=(X+Y)\cap(Z+T)=$$
$$=\{a\in M\colon (\exists r_1,r_2,r_3,r_4\in R)(a=r_1f_2+r_2(f_1-f_2)=r_3f_3+r_4(f_1-f_3))\}=$$
$$=\{a\in M\colon (\exists r_1,r_2,r_3,r_4\in R)(a=r_2f_1+(r_1-r_2)f_2=r_4f_1+(r_3-r_4)f_3)\}=$$
$$=\{a\in M\colon (\exists r_1,r_2,r_3,r_4\in R)(a=r_2f_1+(r_1-r_2)f_2=r_4f_1+(r_3-r_4)f_3\quad\text{and}$$
$$r_2=r_4,\ r_1-r_2=0,\ r_3-r_4=0)\}=\{a\in M\colon (\exists r_2\in R)(a=r_2f_1)\}=\{r_2f_1\colon r_2\in R\}=[f_1].$$

The rest of (4.4) follows similarly via induction. Another elementary computation of the same nature yields

$$q(H^j\colon h\in U)=\{af_1+bf_2+cf_3\colon a,b,c\in R\text{ and }(\exists r\in R)(mn^{j+1}r=kn^ja)\}.$$

Therefore $f_1=1f_1+0f_2+0f_3\in q(H^j\colon h\in U)$ iff $D(mn^{j+1},kn^j)$ holds in R, and a reference to Theorem 3.5 (B) completes the proof.

5. Systems of ring equations. Let u and v be natural numbers and, for $i<v$, let $f_i(y_j\colon j<u)$ be a ring term (i.e., a term in the language of *unitary* rings). Then

$$f_i(y_j\colon j<u)=0,\quad 0\leqq i<v,$$

is called a system of ring equations. This system is said to be solvable in a ring R iff there exist elements r_j, $j<u$, in R such that $f_i(r_j\colon j<u)=0$ for $i<v$.

Lemma 5.1. *For any Horn sentence χ there is a set $\{E_n\colon n<\omega\}$ of systems of ring equations such that*

(i) *for any ring R, χ holds in $\mathbf{L}(R)$ iff there exists an $n<\omega$ such that E_n is solvable in R;*

(ii) *$E_0,E_1,E_2,\ldots$ is a weakening sequence in the sense that, for any $n<\omega$ and any ring R, if E_n is solvable in R then so is E_{n+1}.*

The proof will only be outlined as it is relatively easy but would need a lot of technical preliminaries. First, consider a Mal'tsev condition $(\exists n<\omega)(U_n)$, where $U_0, U_1, U_2, \ldots$ is a weakening sequence of strong Mal'tsev conditions, such that, for any congruence permutable variety $\mathcal{V}$, χ holds in Con $(\mathcal{V})$ iff $(\exists n<\omega)$ (U_n holds in $\mathcal{V}$). The existence of this Mal'tsev condition was proved by Jónsson [13]; a Mal'tsev condition of this kind is explicitly given in [3]. We can easily associate a system E_n of ring equations with each U_n such that, for any ring R, U_n holds in R-**Mod** (which is a congruence permutable variety) iff E_n is solvable in R (cf., e.g., [2, Claim 5.1] or [12, proof of Theorem 2] where analogous or particular cases are handled).

Corollary 5.2. *Let R be the direct product of two rings, R_1 and R_2. Then $\mathbf{L}(R)=\mathbf{L}(R_1)\vee\mathbf{L}(R_2)$ in the lattice of quasivarieties of lattices.*

Proof. We only need to show that an arbitrary Horn sentence χ holds in $\mathbf{L}(R)$ iff it holds in both $\mathbf{L}(R_1)$ and $\mathbf{L}(R_2)$. It is easy to see that a system of ring equations is solvable in R iff it is solvable both in R_1 and R_2. Now if $\mathbf{L}(R_1)\models\chi$ and $\mathbf{L}(R_2)\models\chi$ then, by Lemma 5.1, there are m and k such that the appropriate E_m and E_k are solvable in R_1 and R_2, respectively. Put $n=\max\{m, k\}$. Then E_n is solvable in R_1 and R_2, whence it is solvable in R and $\mathbf{L}(R)\models\chi$. Conversely, if $\mathbf{L}(R)\models\chi$ is assumed then $\mathbf{L}(R_1)\models\chi$ and $\mathbf{L}(R_2)\models\chi$ follows similarly and even more easily.

6. Two results of G. Hutchinson. In this section we will deduce two results of Hutchinson [10] from the results of Sections 3 and 5.

Corollary 6.1 (Hutchinson [10, Proposition 2]). *Assume that R_1 and R_2 are rings and there is a homomorphism of R_1 into R_2 (preserving 1, of course). Then $\mathbf{L}(R_2)\subseteq\mathbf{L}(R_1)$.*

Proof. Let $\varphi\colon R_1\to R_2$ be a ring homomorphism. It suffices to show that any Horn sentence holding in $\mathbf{L}(R_1)$ holds in $\mathbf{L}(R_2)$, too. But this is evident by Lemma 5.1 as φ maps any solution of E_n in R_1 to a solution of E_n in R_2.

Proposition 6.2 (Hutchinson [10]). *Let R_1 and R_2 be rings with the same spectrum $S=S_{R_1}=S_{R_2}$, and assume that either R_1 and R_2 are torsion free or $S(0)$, the characteristic of R_1 and R_2, is a square free (i.e., divisible by p^2 for no prime p) positive number. Then $\mathbf{L}(R_1)=\mathbf{L}(R_2)$.*

Proof. First we prove the statement under the following stronger assumption: either $S(0)$ is a prime or R_1 and R_2 are torsion free. It is sufficient to show that $\mathbf{L}(R_1)$ and $\mathbf{L}(R_2)$ satisfy exactly the same Horn sentences. Therefore it suffices to show that an appropriate construction needed by Theorem 3.5 (B) does not depend (in a sense to be defined later) on the choice of $R\in\{R_1, R_2\}$.

Let $F=\{f_1, f_2, ..., f_t\}$ be a set, let $R\in\{R_1, R_2\}$, and let M be the free R-module generated by F. A submodule C of M will be called *normal*, if it is of the form $[\sum(c_{ij}f_j\colon 1\leqq j\leqq t)\colon i<n_C]$ with a suitable $n_C<\omega$ and integers c_{ij}. The form $[\sum(c_{ij}f_j\colon 1\leqq j\leqq t)\colon i<n_C]$ will be called a *normal form* of C. Note that if only F is fixed then distinct submodules (necessarily over distinct rings) may have identical normal forms. We need

Claim 6.3. Assume that $C, D\in\mathrm{Su}(M)$ are given by the respective normal forms $[\sum(c_{ij}f_j\colon 1\leqq j\leqq t)\colon i<n_C]$ and $[\sum(d_{ij}f_j\colon 1\leqq j\leqq t)\colon i<n_D]$. Then there are normal forms of $C+D$ and $C\cap D$ that depend only on the normal forms $[\sum(c_{ij}f_j\colon 1\leqq j\leqq t)\colon i<n_C]$ and $[\sum(d_{ij}\colon 1\leqq j\leqq t)\colon i<n_D]$ but do not depend on $R\in\{R_1, R_2\}$.

Proof of Claim 6.3. The statement is trivial for $C+D$ as $[\sum(e_{ij}f_j\colon 1\leqq j\leqq t)\colon i<n_C+n_D]$, where $e_{ij}=c_{ij}$ for $i<n_C$ and $e_{ij}=d_{i-n_C,j}$ for $n_C\leqq i<n_C+n_D$, is a normal form of $C+D$. Dealing wiht $C\cap D$, put $n=n_C+n_D$, and let $\mathbf{y}$ and $\mathbf{r}$ stand for n-dimensional column vectors. Then the system of linear equations

$$\sum(c_{ij}y_i\colon i<n_C)-\sum(d_{ij}y_{n_C+i}\colon i<n_D)=0 \quad (1\leqq j\leqq t)$$

can be written of the form $B\mathbf{y}=\mathbf{0}$ for a suitable integer matrix B. It is easy to see that, denoting the entries of $\mathbf{r}$ by r_i,

$$C\cap D=\{\sum(r_i\sum(c_{ij}f_j\colon 1\leqq j\leqq t)\colon i<n_C)\colon \mathbf{r}\in R^n \text{ and } B\mathbf{r}=\mathbf{0}\}.$$

A classical matrix diagonalization method of Frobenius yields that there are integer square matrices A and C of appropriate sizes such that A and C are invertible, their inverses are integer matrices and ABC is a diagonal matrix, i.e., the jth entry of the ith row is 0 whenever $i\neq j$ (cf. Frobenius [7]; this result is quoted with a proof in [12, p. 284 and Appendix]). Denoting $C^{-1}\mathbf{r}$ by $\mathbf{r}'$ and observing that, by the existence of A^{-1}, $B\mathbf{r}=\mathbf{0}$ is equivalent to $AB\mathbf{r}=\mathbf{0}$, we have

$$\{\mathbf{r}\in R^n\colon B\mathbf{r}=\mathbf{0}\}=\{\mathbf{r}\in R^n\colon (ABC)(C^{-1}\mathbf{r})=\mathbf{0}\}=$$

$$=\{\mathbf{r}\in R^n\colon (\exists\mathbf{r}'\in R^n)((ABC)\mathbf{r}'=\mathbf{0} \text{ and } \mathbf{r}=C\mathbf{r}')\}.$$

As ABC is diagonal, $(ABC)\mathbf{r}'=\mathbf{0}$ is equivalent to $g_0r_0'=0, g_1r_1'=0, ..., g_{n-1}r_{n-1}'=0$ where the integers $g_0, g_1, ..., g_{n-1}$ are the diagonal entries of ABC and $r_0', ..., r_{n-1}'$ are the entries of $\mathbf{r}'\in R^n$. For each i, the equation $g_ir_i'=0$ either makes no restriction on r_i' or implies $r_i'=0$. Really, if R is torsion free then $g_i\neq 0$ implies $r_i'=0$; if $S(0)=S_R(0)=p$ is a prime and p does not divide g_i then there is a g' such that $g'g_i\equiv 1 \bmod (p)$ and $g_ir_i'=0$ implies $r_i'=1r_i'=g'g_ir_i'=g'0=0$ while $g_ir_i'=0$ holds for all $r_i'\in R$ when p divides g_i. Put $I=\{i\colon i<n$ and $g_ir_i'=0$ holds for any

$r'_i \in R\}$, and let h_{il}, $0 \leqq i, l < n$, be the entries of the matrix C. Then $r_i = \sum (h_{il} r'_l \colon l < n)$ and we have

$$C \cap D =$$
$$= \{\sum (r_i \sum (c_{ij} f_j \colon 1 \leqq j \leqq t) \colon i < n_C) \colon (\exists \mathbf{r}' \in R^n)(\mathbf{r} = C\mathbf{r}' \text{ and } r'_l = 0 \text{ for all } l \notin I)\} =$$
$$= \{\sum (\sum (h_{il} r'_l \colon l < n) \cdot \sum (c_{ij} f_j \colon 1 \leqq j \leqq t) \colon i < n_C) \colon \mathbf{r}' \in R^n \text{ and } r'_l = 0 \text{ for } l \notin I\} =$$
$$= \{\sum (r'_l \sum (\sum (h_{il} c_{ij} \colon i < n_C) f_j \colon 1 \leqq j \leqq t) \colon l < n) \colon \mathbf{r}' \in R^n \text{ and } r'_l = 0 \text{ for } l \notin I\} =$$
$$= [\sum (\sum (h_{il} c_{ij} \colon i < n_C) f_j \colon 1 \leqq j \leqq t) \colon l \in I],$$

proving Claim 6.3.

Now, returning to the proof of Proposition 6.2, we intend to show that it is possible to choose subsets S_m in Theorem 3.5 (B) so that $F^m = \{f_1, f_2, \dots, f_{t_m}\}$ be the same for $R = R_1$ and $R = R_2$ and, for $x \in U$, X^m be given in a normal form independent of $R \in \{R_1, R_2\}$. This is clearly true for $m = 0$; to start our induction step let us assume that this is true for $m-1$, $m \geqq 1$. Then, by Claim 6.3, $P_m^{m-1} = p_m(X^{m-1} \colon x \in U)$ also has a normal form $[\sum (c_{ij}^{(m)} f_j \colon 1 \leqq j \leqq t_{m-1}) \colon i < n^{(m)}]$ which does not depend on $R \in \{R_1, R_2\}$. Put $S_m := \{s_i \colon i < n^{(m)}\}$ where $s_i = \sum (c_{ij}^{(m)} f_j \colon 1 \leqq j \leqq t_{m-1})$. Then F^m does clearly not depend on $R \in \{R_1, R_2\}$ and, by Claim 6.3,

$$X^m = X^{m-1} + \sum ([v-u] \colon (u, x, v) \in E(G^m) \setminus E(G^{m-1})) =$$
$$= X^{m-1} + \sum (\sum ([v-u] \colon (u, x, v) \in E(\{m\} \times \{s_i\} \times G(q_m))) \colon i < n^{(m)})$$

can be given by a normal form not depending on $R \in \{R_1, R_2\}$. Now a final use of Claim 6.3 yields that, for all m, $q(X^m \colon x \in U)$ can be given by a normal form, say, $[\sum (d_{ij}^{(m)} f_j \colon i \leqq j \leqq t_m) \colon i < k^{(m)}]$ which does not depend on $R \in \{R_1, R_2\}$. Let $\mathbf{y} = (y_0, y_1, \dots, y_{k^{(m)}-1})$, and observe that $f \in q(X^m \colon x \in U)$ iff the following system E_m of *productless* ring equations

$$\sum (d_{i1}^{(m)} y_i \colon i < k^{(m)}) - 1 = 0,$$
$$\sum (d_{ij}^{(m)} y_i \colon i < k^{(m)}) = 0 \quad \text{for} \quad 1 < j \leqq t_m,$$

which does not depend on $R \in \{R_1, R_2\}$, is solvable in R. Based on the afore-mentioned result of Frobenius, it has been shown in [12] (cf. Theorem 2.1 (d) and [12, Theorem 3]) that the solvability of any system of productless ring equations in an arbitrary ring depends only on the spectrum of this ring. But now $S_{R_1} = S = S_{R_2}$, whence E_m is solvable in R_1 iff it is solvable in R_2. Hence Theorem 3.5 (B) proves the proposition under the stronger assumption we considered.

The case $S(0) = 1$ being trivial, consider the case $S(0) = p_0 p_1 \dots p_n$ where $p_0, p_1, \dots, p_n$ are distinct primes. It is known that, for $R \in \{R_1, R_2\}$, R is isomorphic to a direct product $\prod_{i \leqq n} R^{(i)}$ where $S_{R^{(i)}}(0) = p_i$, $i \leqq n$ (cf., e.g., McCoy [14, Theorem 28]). By Corollary 5.2, $\mathbf{L}(R)=\bigvee_{i\leqq n}\mathbf{L}(R^{(i)})$, whence Proposition 6.2 follows from its special case we have already proved.

7. Two sufficient conditions for regularity. Consider a Horn sentence χ of the form (2.2).

Proposition 7.1. *If all q_i, $0\leqq i\leqq t$, are join-free then χ is regular.*

Proof. We will use Theorem 3.5 (B) with $S_m:=P_m^{m-1}$. As $G(q_m)$ has no inner vertex, $F^m=F^{m-1}=\ldots=F^0$, if $x\in U$ occurs in q_m then $X^m=X^{m-1}+P_m^{m-1}$, and $X^m=X^{m-1}$ if $x\in U$ does not occur in q_m. Hence there are lattice terms q_m' such that $q(X^m\colon x\in U)=q_m'(X^0\colon x\in U)$, $0\leqq m$, and these q_m' do not depend on the ring in question. By Theorem 3.5 (B), $\mathbf{L}(R)\models\chi$ is equivalent to $(\exists m)\big(f_1\in q_m'(X^0\colon x\in U)\big)$.

Now let us fix a $y\in U$ and consider the Horn sentence χ_k: $y\leqq y\Rightarrow p\leqq q_k'$, $k\geqq 0$. If we apply Theorem 3.5 (B) to χ_k with $S_m=\emptyset$, $1\leqq m$, then $X^m=X^{m-1}=\ldots=X^0$ for every $x\in U$. Hence $f_1\in q_k'(X^0\colon x\in U)$ is equivalent to $\mathbf{L}(R)\models\chi_k$. But χ_k, being modulo lattice theory equivalent to the lattice identity $p\leqq q_k'$, is regular by Theorem 2.1 (a). We have seen that $\mathbf{L}(R)\models\chi$ is equivalent to $(\exists m)\big(\mathbf{L}(R)\models\chi_m\big)$, whence the regularity of χ_m completes the proof.

Note that Proposition 7.1 applies for $\chi(m,n,k)$ occurring in Theorem 4.1. We say that χ satisfies the Whitman condition (W) if the finitely presented lattice $FL(U;p_0\leqq q_0, p_1\leqq q_1,\ldots,p_t\leqq q_t)$ satisfies (W) (cf. [5]).

Proposition 7.2. *If χ satisfies (W) then χ is regular.*

Proof. By [5, Corollary 5.3] there are lattice identites $\varkappa_m$, $m<\omega$, such that, for any n-permutable variety $\mathscr{V}$, $\mathrm{Con}(\mathscr{V})\models\chi$ iff $(\exists m)(\mathrm{Con}(\mathscr{V})\models\varkappa_m)$. In particular, $\mathbf{L}(R)\models\chi$ iff $\mathrm{Con}(R\text{-}\mathbf{Mod})\models\chi$ iff $(\exists m)(\mathrm{Con}(R\text{-}\mathbf{Mod})\models\varkappa_m)$ iff $(\exists m)\big(\mathbf{L}(R)\models\varkappa_m\big)$, whence the regularity of the lattice identites $\varkappa_m$ (cf. Theorem 2.1 (a)) completes the proof.

8. Proof of Theorem 4.2. With the notations of Section 2, let us recall

Claim 8.1. (Hutchinson [12, Proposition 4 and the proof of Theorem 5] or, more explicitly, [2, Proposition 6.2]). If $S_1, S_2\in\mathscr{L}_S$ and $S_1\leqq S_2$ then R_{S_1} is a homomorphic image of R_{S_2}.

Given a spectrum $S\in\mathscr{L}_S$, let $\{p\colon p\in P$ and $S(p)<\omega\}$ be denoted by $T(S)$. Let S be called *cofinite* iff $T(S)$ is finite. Note that $S(0)=0$ for any cofinite $S\in\mathscr{L}_S$. If S is an arbitrary spectrum and H is a finite subset of P then the spectrum $S[H]$ defined by $S[H](0)=0$, $S[H](p)=S(p)$ for $p\in H$ and $S[H](p)=\omega$ for $p\in P\setminus H$ is cofinite, and we have $S\leqq S[H]$ and $T(S[H])\subseteq H$.

Now let us fix a regular Horn sentence χ which holds in $\mathbf{L}(R^*)$ for some ring R^* with $S_{R^*}(0)=0$. Put $S^*=S_{R^*}$. Since χ is regular, it holds in $\mathbf{L}(R_{S^*})$ by Theorem 2.1 (c). Let S^0 denote the zero spectrum, i.e., $S^0(x)=0$ for all $x\in\{0\}\cup P$, and put $R^0=R_{S^0}$. (Note that S^0 is not the smallest element of $\mathscr{L}$.) Then $S^0\leqq S^*$, whence, by Corollary 6.1 and Claim 8.1, χ holds in $L(R^0)$.

Now consider the system of ring equations $E_0, E_1, E_2, \ldots$ associated with χ by Lemma 5.1.

Claim 8.2. Let $S\in\mathscr{L}_S$ with $S(0)=0$ and let n be a non-negative integer. If E_n is solvable in R_S then there is a finite subset H of P such that E_n is solvable in $R_{S[H]}$.

Proof. Let E_n consist of the ring equations $f_i(y_j\colon j<u)=0$, $i<v$, and assume that $f_i(a_j+J_S\colon j<u)=0+J_S$, i.e., $f_i(a_j\colon j<u)\in J_S$, $i<v$, for certain elements $a_i\in FR(\{x_p\colon p\in P\})$, $i<v$ (cf. Section 2). As we have only finitely many $f_i(a_j\colon j<u)$, there is a finite subset A of $\{p^{S(p)}(px_p-1)\colon p\in P \text{ and } S(p)<\omega\}$ such that all the $f_i(a_j\colon j<u)$, $i<v$, belong to the ideal generated by A. Put $H=\{p\colon p\in P,\ S(p)<\omega$ and $p^{S(p)}(px_p-1)\in A\}$. Then H is finite, and $A\subseteqq J_{S[H]}$ yields $f_i(a_j\colon j<u)\in J_{S[H]}$ for all $i<v$. Hence the system of $a_j+J_{S[H]}$, $j<u$, is a solution of E_n in $R_{S[H]}$.

Since $E_0, E_1, E_2, \ldots$ is a weakening sequence, the first n_0 of its members can be omitted without the loss of generality, for any $n_0<\omega$. Therefore, by Lemma 5.1, we may assume that E_0 is solvable in R^0. Hence, by Claim 8.2, we can fix a cofinite spectrum S' such that E_0, and therefore every E_n, is solvable in $R_{S'}$. (Indeed, let $S'=S^0[H]$ for an appropriate $H\subseteqq P$.)

For $j<\omega$, let $U_j:=\{S_R\colon R$ is a ring and E_j is solvable in $R\}$. Then $U_j\subseteqq\mathscr{L}_S$. For each $S\in U_j$ choose a ring $B_{j,S}$ such that E_j is solvable in $B_{j,S}$ and the spectrum of $B_{j,S}$ is S. Put $A_j:=\prod(B_{j,S}\colon S\in U_j)$, the direct product of $B_{j,S}$, $S\in U_j$, and let $S_j:=\bigvee(S\colon S\in U_j)$.

Claim 8.3. The spectrum of A_j is S_j, S_j is cofinite, E_j is solvable in A_j, and, for any ring R, if E_j is solvable in R then $S_R\leqq S_j$.

Proof. E_j is clearly solvable in A_j as it is solvable in all the direct factors of A_j. Similarly, a divisibility condition $D(m, n)$, which is a particular ring equation, holds in A_j iff $D(m, n)$ holds in every $B_{j,S}$, $S\in U_j$. As $S'\in U_j$ and $S'(0)=0$, $S_j(0)=0$ and the characteristic of $A_j=B_{j,S'}\times\prod(B_{j,S}\colon S\in U_j\setminus\{S'\})$ is also 0. Further, by Theorem 2.1 (d), we have

$$\min\{i\colon A_j\models D(p^{i+1}, p^i)\} = \min\bigcap(\{i\colon B_{j,S}\models D(p^{i+1}, p^i)\}\colon S\in U_j) =$$

$$= \min\bigcap(\{i\colon i = \exp(p^i, p) \geqq S(p)\}\colon S\in U_j) = \sup\{S(p)\colon S\in U_j\} = S_j(p)$$

for any $p\in P$. Consequently, S_j is the spectrum of A_j. From $S'\leqq S_j$ we obtain $T(S')\supseteqq T(S_j)$, whence S_j is cofinite. Finally, if E_j is solvable in a ring R then $S_R\in U_j$ yields that $S_R\leqq\bigvee(S\colon S\in U_j)=S_j$.

Now let $I:=\{S\in\mathscr{L}_S\colon \chi$ holds in $\mathbf{L}(R_S)\}$, and let $(S_j]$ denote $\{S\in\mathscr{L}_S\colon S\leqq S_j\}$, the principal ideal of $\mathscr{L}_S$ generated by S_j.

Claim 8.4. $I=\bigcup((S_j]\colon j<\omega)$.

Proof. If $S\in I$ then, by Lemma 5.1, there is a $j<\omega$ such that E_j is solvable in R_S. Hence $S=S_{R_S}\in(S_j]$ by Claim 8.3. Conversely, assume that $S\in(S_j]$ for some $j<\omega$. By Lemma 5.1 and Claim 8.3, χ holds in $\mathbf{L}(A_j)$. By Theorem 2.1 (c) and Claim 8.3, A_j and R_{S_j} have the same spectrum S_j. The regularity of χ yields that χ holds in $\mathbf{L}(R_{S_j})$, too. Now $S\in I$ follows from $S\leqq S_j$, Claim 8.1 and Corollary 6.1.

Now we obtain $S_0\leqq S_1\leqq S_2\leqq\ldots$ from the fact that $E_0, E_1, E_2, \ldots$ is a weakening sequence. Hence $T(S_0)\supseteqq T(S_1)\supseteqq T(S_2)\supseteqq\ldots$. Since $T(S_0)$ is finite, so is $H:=\bigcap(T(S_j)\colon j<\omega)$. Put $S:=\bigvee(S_j\colon j<\omega)$, then $T(S)\subseteqq H$. Define m_χ, n_χ and k_χ as follows:

$$m_\chi:=\prod(p^{S(p)+1}\colon p\in T(S)),\quad n_\chi:=\prod(p\colon p\in H\setminus T(S))$$

and

$$k_\chi:=\prod(p^{S(p)}\colon p\in T(S)).$$

Then m_χ, n_χ and k_χ are positive integers.

Assume that $D(m_\chi n_\chi^{i+1}, k_\chi n_\chi^i)$ holds in a ring R for some $i<\omega$. Then, by Theorem 2.1 (d), we have $S(p)=\exp(k_\chi n_\chi^i, p)\geqq S_R(p)$ for $p\in T(S)$ and $S(p)=$ $=\omega>i=\exp(k_\chi n_\chi^i, p)\geqq S_R(p)$ for $p\in H\setminus T(S)$. For $p\in H$, $S(p)$ is the limit of the increasing sequence $S_0(p), S_1(p), S_2(p), \ldots$, whence the finiteness of H yields the existence of a $j<\omega$ such that $T(S_j)=H$ and $S_R(p)\leqq S_j(p)$ for all $p\in H$. Then $S_R\leqq S_j$, and Claim 8.4 yields that $S_R\in I$. I.e., χ holds in $\mathbf{L}(R_{S_R})$. Since R_{S_R} and R have the same spectrum and χ is regular, χ holds in $\mathbf{L}(R)$.

Conversely, assume that R is a ring and χ holds in $\mathbf{L}(R)$. As R_{S_R} and R have the same spectrum and χ is regular, $S_R\in I$. By Claim 8.4, there is a $j<\omega$ such that $T(S_j)=H$ and $S_R\leqq S_j\leqq S$. Put $i=\max\{S_j(p)\colon p\in H\setminus T(S)\}$, then i is a nonnegative integer. (Here $\max\emptyset=0$.) For $p\in T(S)$, $\exp(k_\chi n_\chi^i, p)=S(p)\geqq S_R(p)$ while, for $p\in H\setminus T(S)$, $\exp(k_\chi n_\chi^i, p)=i\geqq S_j(p)\geqq S_R(p)$. Hence, by Theorem 2.1 (d), $D(m_\chi n_\chi^{i+1}, k_\chi n_\chi^i)$ holds in R. This completes the proof of Theorem 4.2.

9. On a problem of Jónsson. In this section we will give a negative answer to (1.4), the afore-mentioned problem of Jónsson [13]. Let $n\geqq2$ be an integer, and consider $\chi(0, n, 1)$ from Theorem 4.1. Then we have

Proposition 9.1. *There is no strong Mal'tsev condition U such that, for any congruence permutable variety $\mathcal{V}$, $\chi(0, n, 1)$ holds in* $\mathrm{Con}(\mathcal{V})$ *iff U holds in $\mathcal{V}$.*

Proof. Assume the contrary, and let E be a system of ring equations such that, for any ring R, U holds in R-**Mod** iff E is solvable in R (cf. the proof of Lemma 5.1). Since $D(0, n^i)$ holds in $\mathbf{Z}_{n^i}$, $\chi(0, n, 1)$ holds in $\mathbf{L}(\mathbf{Z}_{n^i})$ by Theorem 4.1, and we infer that E is solvable in $\mathbf{Z}_{n^i}$, $i<\omega$. Therefore E is solvable in the direct product $R=\prod(\mathbf{Z}_{n^i}\colon i<\omega)$, whence $\chi(0, n, 1)$ holds in $\mathbf{L}(R)$. It follows from Theorem 4.1 that there is a $j<\omega$ such that $D(0, n^j)$ holds in R. Consequently, $D(0, n^j)$ holds in every direct factor $\mathbf{Z}_{n^i}$ of R. In particular, $D(0, n^j)$ holds in $\mathbf{Z}_{n^{j+1}}$, which is a contradiction.

References

[1] G. Birkhoff, *Lattice Theory*, 3rd ed., Colloquium Publications 25, Amer. Math. Soc. (Providence, R. I., 1967).

[2] G. Czédli, On properties of rings that can be characterized by infinite lattice identities, *Studia Sci. Math. Hungar.*, **16** (1981), 45—60.

[3] G. Czédli, Mal'cev conditions for Horn sentences with congruence permutability, *Acta Math. Hungar.*, **44** (1984), 115—124.

[4] G. Czédli, Characterizing universal algebras by lattice identities, Thesis for the Candidate's Degree, 1983. (Hungarian)

[5] G. Czédli and A. Day, Horn sentences with (W) and weak Mal'cev conditions, *Algebra Universalis*, **19** (1984), 217—230.

[6] G. Czédli and G. Hutchinson, An irregular Horn sentence in submodule lattices, *Acta Sci. Math.*, **51** (1987), 35—38.

[7] G. Frobenius, Theorie der linearen Formen mit ganzen Coefficienten, *J. Reine Angew. Math.*, **86** (1879), 146—208.

[8] C. Herrmann und A. P. Huhn, Zum Begriff der Charakteristik modularer Verbände, *Math. Z.*, **144** (1975), 185—194.

[9] C. Herrmann and W. Poguntke, Axiomatic classes of lattices of normal subgroups, Preprint, 12, Technische Hochschule Darmstadt, 1972.

[10] G. Hutchinson, On classes of lattices representable by modules, in: *Proc. Univ. Houston Lattice Theory Conf.* (Houston, 1973); pp. 69—94.

[11] G. Hutchinson, On the representation of lattices by modules, *Trans. Amer. Math. Soc.*, **209** (1975), 311—351.

[11B] G. Hutchinson, Exact embedding functors between categories of modules, *J. Pure Appl. Algebra*, **25** (1982), 807—111.

[12] G. Hutchinson and G. Czédli, A test for identities satisfied in lattices of submodules, *Algebra Universalis*, **8** (1978), 269—309.

[13] B. Jónnson, Congruence varieties, *Algebra Universalis*, **10** (1980), 355—394.

[14] N. H. McCoy, *Rings and Ideals*, Charus Math. Monographs, 8, Math. Assoc. Amer., Waverly Press (Baltimore, 1948).

JATE BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math., 51 (1987), 35—38

An irregular Horn sentence in submodule lattices

GÁBOR CZÉDLI*) and GEORGE HUTCHINSON

Dedicated to the memory of András P. Huhn

For a ring R, always with 1, a lattice is said to be representable by R-modules if it is embeddable in the lattice of submodules of some unital left R-module. Let $\mathbf{L}(R)$ denote the class of lattices representable by R-modules. Then $\mathbf{L}(R)$ is known to be a quasivariety, i.e., to be axiomatizable by (universal) Horn sentences (cf. e.g., [5]). Let $\mathbf{HL}(R)$ denote the lattice variety generated by $\mathbf{L}(R)$. A Horn sentence χ is called ***irregular*** (cf. [1]) if there are rings R_1 and R_2 such that $\mathbf{HL}(R_1)=\mathbf{HL}(R_2)$ and χ holds in $\mathbf{L}(R_1)$ but χ does not hold in $\mathbf{L}(R_2)$. Although the existence of irregular Horn sentences follows from [4, p. 92], no concrete irregular Horn sentence was known previously. The aim of the present note is to give an irregular Horn sentence $\hat{\chi}$. This $\hat{\chi}$ was found by applying the techniques of [1] and generalizing the methods of HERRMANN and HUHN [3] and [8]. Note that regular Horn sentences are much easier to handle, cf. [1].

Consider the following lattice terms on the set $U=\{x, y, z, t\}$ of variables:

$$\begin{aligned} p &= (x+y)(z+t), & h_0 &= (x+z)(y+t),\\ h_1 &= (x+t)(y+z), & h_2 &= (x+t)(p+h_0),\\ h_3 &= (y+t)(h_1+p), & p_0 &= (h_2+z)y,\\ q_0 &= x+z+h_3, & q &= p_0+x, \end{aligned}$$

and let $\hat{\chi}$ be the Horn sentence

$$p_0 \leqq q_0 \Rightarrow p \leqq q.$$

Theorem. *$\hat{\chi}$ is irregular.*

Received April 29, 1986.

*) Research partially supported by Hungarian National Foundation for Scientific Research grant no. 1813.

Proof. Let $\mathbf{Z}_4$ stand for the factor ring of the ring of integers modulo 4. Let I_1 and I_2 denote the ideals of $\mathbf{Z}_4[x]$ generated by $\{x^2-2, 2x\}$ and $\{x^2, 2x\}$, respectively. The rings $R_1=\mathbf{Z}_4[x]/I_1$ and $R_2=\mathbf{Z}_4[x]/I_2$ consist of eight elements. With the notations $a=x+I_1$ and $b=x+I_2$, we have

$$R_1 = \{i+ja\colon 0 \leqq i \leqq 3,\ 0 \leqq j \leqq 1\} \quad \text{and} \quad R_2 = \{i+jb\colon 0 \leqq i \leqq 3,\ 0 \leqq j \leqq 1\}.$$

Moreover, the bijection $\varphi\colon R_1 \to R_2,\ i+ja \mapsto i+jb$ preserves the unit element and the additive structure. Therefore, $\mathbf{HL}(R_1)=\mathbf{HL}(R_2)$ (cf. [8, Prop. 3]). So, it suffices to show that χ holds in $\mathbf{L}(R_1)$ but does not hold in $\mathbf{L}(R_2)$.

As Theorem 3.5 of [1] will be our main tool, we adopt the notations preceeding the theorem in [1, § 3]. First, by [1, Thm. 3.5 (A)], we prove that χ holds in $\mathbf{L}(R_1)$. Now $p_j=p_0$ and $q_j=q_0$ for $j\geqq 1$, and $F^0=\{f_1, f_2, f_3\}$ according to Figure 1. We have $X^0=[f_2]$, $Y^0=[f_1-f_2]$, $Z^0=[f_3]$ and $T^0=[f_1-f_3]$. Denoting $k(C^m\colon c\in U)$ by K^m for $m\geqq 0$ and $k\in\{p, q, p_0, q_0, h_0, h_1, h_2, h_3\}$, an elementary calculation in $\mathrm{Su}(M^0)$ shows that $P^0=[f_1]$, $H_0^0=[f_2-f_3]$, $H_2^0=[f_1+f_2-f_3]$ and $P_1^0= =P_0^0=\{r(f_1-f_2)\colon r\in R_1 \text{ and } 2r=0\}$. Since $2a=0$, we may choose $S_1=\{a(f_1-f_2)\}$. Let $F^1=\{f_1, f_2, f_3, e_1, e_2, \ldots, e_8\}$ according to Figure 2.

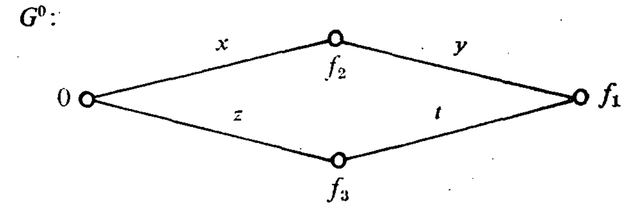


Figure 1

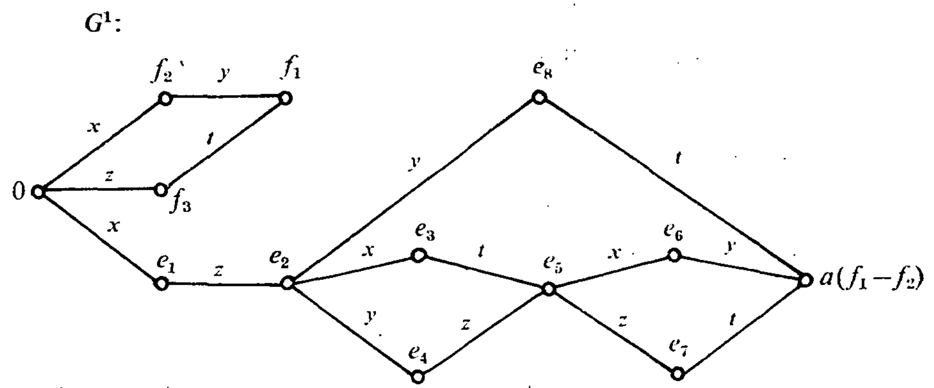


Figure 2

We obtain the following formulas, each of them an easy consequence of the previous ones or Figure 2.

$$X^1 = [f_2, e_1, e_2-e_3, e_5-e_6],$$
$$Y^1 = [f_1-f_2, e_2-e_4, e_2-e_8, e_6],$$
$$Z^1 = [f_3, e_1-e_2, e_4-e_5, e_5-e_7],$$
$$T^1 = [f_1-f_3, e_3-e_5, e_7-e_8, a(f_1-f_2)-e_7],$$
$$P^1 \supseteq [f_1, e_3-e_4, af_2+e_5],$$
$$H_0^1 \supseteq [f_2-f_3, e_3-e_5+e_6, e_4-e_6],$$
$$H_2^1 \supseteq [f_1+f_2-f_3, e_3-e_6, a(f_1-f_3)+e_3-2e_5+e_6].$$

Since $a^2=2$ and $2a=0$,

$$f_1-f_2 = -(f_1+f_2-f_3)+a(e_3-e_6)+a\big(a(f_1-f_3)+e_3-2e_5+e_6\big)+f_3 \in H_2^1+Z^1.$$

Therefore, we have $f_1=(f_1-f_2)+f_2 \in P_0^1+X^1=Q^1=q(C^1\colon c\in U)$. Hence χ holds in $\mathbf{L}(R_1)$ by [1, Thm. 3.5 (A)].

Now observe that I_2 is included in the ideal I of $\mathbf{Z}_4[x]$ generated by x, whence $\mathbf{Z}_4 \approx \mathbf{Z}_4[x]/I$ is a homomorphic image of R_2. Therefore, if χ held in $\mathbf{L}(R_2)$, it would also hold in $\mathbf{L}(\mathbf{Z}_4)$ by [4, Prop. 2] (or by [1, Cor. 6.1]). Hence it suffices to show that χ does not hold in $\mathbf{L}(\mathbf{Z}_4)$. As suggested by [1, Thm. 3.5 (B)], we let $x=\mathbf{Z}_4 f_2$, $y=\mathbf{Z}_4(f_1-f_2)$, $z=\mathbf{Z}_4 f_3$ and $t=\mathbf{Z}_4(f_1-f_3)$ in a free $\mathbf{Z}_4$-module with three generators f_1, f_2 and f_3. Calculation shows that $p_0=\mathbf{Z}_4(2f_1+2f_2)$, $q_0=\mathbf{Z}_4 2f_1+\mathbf{Z}_4 f_2+\mathbf{Z}_4 f_3$, $p=\mathbf{Z}_4 f_1$ and $q=\mathbf{Z}_4 2f_1+\mathbf{Z}_4 f_2$. Therefore, χ fails in $\mathbf{L}(\mathbf{Z}_4)$, proving the theorem.

In [4, p. 92], it was shown that no R_1-module is a free $\mathbf{Z}_4$-module (a direct sum of cyclic groups of order 4). This is the key property allowing construction of an irregular Horn sentence, as observed below.

Let S denote $\mathbf{Z}/p^k\mathbf{Z}$, the ring of integers moduIo p^k for p prime and $k \geqq 2$. We show that $\mathbf{L}(R)=\mathbf{L}(S)$ if and only if R has characteristic p^k and some (nontrivial) R-module M is free as an S-module (that is, M is a direct sum of cyclic groups of order p^k).

Supposing $\mathbf{L}(R)=\mathbf{L}(S)$, R has characteristic p^k (cf. [1, Thm. 2.1]). By [6, Thm. 1, p. 108], there is an exact embedding functor F from S-Mod into R-Mod. For $n\cdot f=f+\ldots+f$ (n times), we see that $\langle p\cdot 1_A, p^{k-1}\cdot 1_A\rangle$ is exact in R-Mod for $A=F({}_SS)\neq 0$. Since A is a direct sum of cyclic groups, each with order dividing p^k (PRÜFER, see [2, Thm. 17.2, p. 88], it follows that A is free as an S-module.

For the converse, note that an R-module M which is free as an S-module can be regarded as a bimodule ${}_RM_S$, which induces an exact embedding S-Mod$\to R$-Mod by the tensor product functor ${}_RM_S \otimes_S$ —, yielding $\mathbf{L}(S)\subseteqq\mathbf{L}(R)$ by [6, Thm. 1,

p. 108]. Since R has characteristic p^k, there is a ring homomorphism $S\to R$. Then $\mathbf{L}(R)=\mathbf{L}(S)$ (cf. [1, Cor. 6.1]).

This result can be regarded as a corollary of the ring theory result proved in [7]: If R and S are nontrivial rings with S left artinian, then there exists an exact embedding functor S-Mod$\to R$-Mod if and only if there exists a nontrivial bimodule ${}_RA_S$ such that A_S is a free right S-module.

References

[1] G. Czédli, Horn sentences in submodule lattices, *Acta Sci. Math.*, **51** (1987), 17—33.

[2] L. Fuchs, *Infinite Abelian Groups*, Vol. I, Academic Press (New York, 1970).

[3] C. Herrmann and A. P. Huhn, Zum Begriff der Charakteristik modularer Verbande, *Math. Z.*, **144** (1975), 185—194.

[4] G. Hutchinson, On classes of lattices representable by modules, in: *Proc. Univ. of Houston Lattice Theory Conference* (Houston, 1973); pp. 69—94.

[5] G. Hutchinson, On the representation of lattices by modules, *Trans. Amer. Math. Soc.*, **209** (1975), 311—351.

[6] G. Hutchinson, Exact embedding functors between categories of modules, *J. Pure Appl. Algebra*, **25** (1982), 107—111.

[7] G. Hutchinson, Addendum to exact embedding functors between categories of modules, *J. Pure Appl. Algebra*, to appear.

[8] G. Hutchinson and G. Czédli, A test for identities satisfied in lattices of submodules, *Algebra Universalis*, **8** (1978), 269—309.

(G. C.)
JATE BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

(G. H.)
NATIONAL INSTITUTES OF HEALTH
BLDG. 12A, ROOM 3045
BETHESDA, MD 20892, USA

Acta Sci. Math., 51 (1987), 39—55

Quasi-identities, Mal'cev conditions and congruence regularity

B. A. DAVEY, K. R. MILES and V. J. SCHUMANN

Dedicated to the memory of András P. Huhn

This work grew out of our desire to present a uniform approach to the various forms of congruence regularity which have been studied in the literature. We were particularly interested in the result of GRÄTZER [8] that if every algebra in a variety $\mathcal{V}$ contains an element a such that $[a]\alpha=[a]\beta$ implies $\alpha=\beta$ for all congruences α, β on A, then the element a may be chosen from any subalgebra of A. We also wished to study the concept of subregularity introduced by TIMM [15]: an algebra A is *subregular* if for all subalgebras $B\leqq A$ and all congruences α, β on A we have $\alpha=\beta$ whenever $[b]\alpha=[b]\beta$ for all $b\in B$. In particular we wanted a characterization of subregularity via simple identities and quasi-identities similar to those for regularity due to WILLE [16] and CSÁKÁNY [2]. These two topics turned out to be quite closely related (Theorem 2.3).

In the first section the various types of regularity are defined and their local properties are investigated. In particular, we give characterizations in terms of principal congruences similar to those for regularity and weak regularity give in HASHIMOTO [12] and GRÄTZER [8] (Lemma 1.3). We also apply GUMM's Shifting Principle [9] to give local proofs of congruence modularity where possible (Theorem 1.4).

The global relationships between the various forms of regularity are studied in Section 2. The section begins with a general translation principle for converting a Hashimoto-type principal-congruence property into a quasi-identity (Theorem 2.1) which is then applied to yield quasi-identity characterizations for each of the forms of regularity (Theorems 2.2, 2.3, 2.4, 2.5).

The third section contains a general consideration of the relationships between quasi-identities, identities, congruence modularity and n-permutability. Several ways of translating quasi-identities into identities are given (Theorems 3.4, 3.5). We

Received June 16, 1986.

also describe a large class of quasi-identities which imply both congruence modularity and n-permutability for some n (Theorem 3.9).

In Section 4 we see that the results of Sections 2 and 3 may be combined to yield identities characterizing each of the forms of regularity and show that, with one exception, each implies congruence modularity and n-permutability for some n.

Our notation and terminology are fairly standard. Note in particular that the lattice of congruences of an algebra A is denoted by Con A with least element 0, the n-generated free algebra in a variety $\mathcal{V}$ is denoted by $F\mathcal{V}(n)$ and by a *constant term* we mean a nullary or constant unary term.

1. Definitions and local relationships. In this section we introduce various degrees of regularity and study these at the local level. A Hashimoto-type principal-congruence characterization is given for each and, where possible, a local proof of congruence modularity is obtained via H. P. Gumm's Shifting Principle.

An algebra A is *regular with respect to* $a_1, \ldots, a_n \in A$ if for all $\alpha, \beta \in \mathrm{Con}\, A$ we have

$$\left(\mathop{\&}_{i=1}^{n} [a_i]\alpha = [a_i]\beta\right) \Rightarrow \alpha = \beta.$$

R: A is *regular* if it is regular with respect to a for each $a \in A$.

R_n: A is *n-regular* if there exist $a_1, \ldots, a_n \in A$ such that A is regular with respect to $a_1, \ldots, a_n$.

SR: A is *subregular* if it is regular with respect to each of its subalgebras, that is, for each $B \leqq A$ and all $\alpha, \beta \in \mathrm{Con}\, A$ we have

$$\left(\mathop{\&}_{b \in B} [b]\alpha = [b]\beta\right) \Rightarrow \alpha = \beta.$$

SR_n: A is *n-subregular* if for all $B \leqq A$ there exist $b_1, \ldots, b_n \in B$ such that A is regular with respect to $b_1, \ldots, b_n$.

Note that 1-regularity is usually referred to as *weak regularity*. Some authors have insisted that the elements $a_1, \ldots, a_n$ in the definition of n-regularity be constant terms: if there are constant terms $o_1, \ldots, o_n$ such that A is regular with respect to $o_1, \ldots, o_n$ then we shall say that A satisfies $\mathsf{R}(o_1, \ldots, o_n)$. We say that a class $\mathcal{V}$ of algebras is *regular* (respectively, *subregular*, etc.) if every algebra in $\mathcal{V}$ is regular (respectively, subregular, etc.).

In Timm [15] it is pointed out that the algebra $\langle \mathbf{N}; s\rangle$, where s is the successor function, is subregular and it is easily seen that it is not n-regular for any $n \in \mathbf{N}$. The non-zero congruences on $\langle \mathbf{N}; s\rangle$ are all of the form $\Theta(m, k)$ for some $m, k \in \mathbf{N}$ where

$$x\Theta(m,k)\,y \;\Leftrightarrow\; x = y < m \quad \text{or} \quad (x, y \geqq m \;\&\; x \equiv y \pmod{k}).$$

Since $\Theta(m, k) \leqq \Theta(n, l) \Leftrightarrow n \leqq m$ and $l|k$, we have

$$\mathrm{Con}\,\langle \mathbf{N}; s\rangle \cong 1 \oplus [\langle \mathbf{N}; \leqq\rangle^d \times \langle \mathbf{N}; |\rangle^d].$$

Note that $\langle \mathbf{N}; s\rangle$ is congruence-distributive.

The implications in the diagram below are trivial:

$$\begin{array}{ccccccc}
 & & \mathsf{R}_1 & \longrightarrow & \mathsf{R}_2 & \to \dots & \\
 & & \uparrow & & \uparrow & & \\
\mathsf{R} & \to & \mathsf{SR}_1 & \longrightarrow & \mathsf{SR}_2 & \to \dots \to & \mathsf{SR} \\
 & & \uparrow & & \uparrow & & \\
 & & \mathsf{R}(o_1) & \to & \mathsf{R}(o_1, o_2) & \to \dots &
\end{array}$$

In the presence of a one-element subalgebra many of these relations collapse.

1.1. Lemma. (i) *If A has a one-element subalgebra, then on A we have* $\mathsf{SR} \to \mathsf{R}_1$.

(ii) *If there is a constant term o such that $\{o\} \leqq A$, then on A we have* $\mathsf{SR} \to \mathsf{R}(o)$.

The following characterizations using set inclusion rather than equality are often useful. If $\alpha \in \mathrm{Con}\,A$ and $B \leqq A$ then $\alpha \restriction B$ denotes the restriction of α to B and $[B]\alpha$ denotes the union over $b \in B$ of the α-blocks $[b]\alpha$.

1.2. Lemma. (i) *A is regular if and only if*

$$(\forall a \in A)(\forall \alpha, \beta \in \mathrm{Con}\,A)[[a]\alpha \subseteqq [a]\beta \Rightarrow \alpha \subseteqq \beta].$$

(ii) *A is n-regular if and only if*

$$(\exists a_1, \dots, a_n \in A)(\forall \alpha, \beta \in \mathrm{Con}\,A)\left[\left(\mathop{\&}_{i=1}^{n}[a_i]\alpha \subseteqq [a_i]\beta\right) \Rightarrow \alpha \subseteqq \beta\right].$$

(iii) *A is n-subregular if and only if*

$$(\forall B \leqq A)(\exists b_1, \dots, b_n \in B)(\forall \alpha, \beta \in \mathrm{Con}\,A)\left[\left(\mathop{\&}_{i=1}^{n}[b_i]\alpha \subseteqq [b_i]\beta\right) \Rightarrow \alpha \subseteqq \beta\right].$$

(iv) *The following are equivalent:*

(a) *A is subregular;*

(b) $(\forall B \leqq A)(\forall \alpha, \beta \in \mathrm{Con}\,A)\left[\left(\mathop{\&}_{b \in B}[b]\alpha \subseteqq [b]\beta\right) \Rightarrow \alpha \subseteqq \beta\right]$;

(c) $(\forall B \leqq A)(\forall \alpha, \beta \in \mathrm{Con}\,A)[(\alpha \restriction B \subseteqq \beta \restriction B \,\&\, [B]\alpha = B) \Rightarrow \alpha \subseteqq \beta]$.

Proof. These proofs are trivial once we observe that $[a]\alpha \subseteqq [a]\beta$ implies $[a]\alpha = [a](\alpha \wedge \beta)$.

The version of subregularity given in 1.2 (iv) (c) has been useful in the study of injectivity: see DAVEY and KOVÁCS [3].

We now give the Hashimoto-type principal-congruence characterizations of the various forms of regularity. The subalgebra generated by $a \in A$ is denoted by $\langle a\rangle$.

1.3. Lemma. (i) *A is regular with respect to $a_1, \ldots, a_n \in A$ if and only if for all $b, c \in A$ there exist $d_{i1}, \ldots, d_{im} \in A$ such that*

$$\Theta(b, c) = \bigvee_{i=1}^{n} \bigvee_{j=1}^{m} \Theta(a_i, d_{ij}).$$

(ii) $A \models \mathsf{R} \Leftrightarrow (\forall a \in A)(\forall b, c \in A)(\exists d_1, \ldots, d_m \in A) \quad \Theta(b, c) = \bigvee_{j=1}^{m} \Theta(a, d_j).$

(iii) $A \models \mathsf{R}_n \Leftrightarrow (\exists a_1, \ldots, a_n \in A)(\forall b, c \in A)(\exists d_{i1}, \ldots, d_{im} \in A)$

$$\Theta(b, c) = \bigvee_{i=1}^{n} \bigvee_{j=1}^{m} \Theta(a_i, d_{ij}).$$

(iv) $A \models \mathsf{SR} \Leftrightarrow (\forall a \in A)(\forall b, c \in A)(\exists a_1, \ldots, a_n \in \langle a \rangle)(\exists d_1, \ldots, d_n \in A)$

$$\Theta(b, c) = \bigvee_{i=1}^{n} \Theta(a_i, d_i).$$

(v) $A \models \mathsf{SR}_n \Leftrightarrow (\forall a \in A)(\exists a_1, \ldots, a_n \in \langle a \rangle)(\forall b, c \in A)(\exists d_{i1}, \ldots, d_{im} \in A)$

$$\Theta(b, c) = \bigvee_{i=1}^{n} \bigvee_{j=1}^{m} \Theta(a_i, d_{ij}).$$

Proof. (ii) is due to Hashimoto [12] and Grätzer [8]. As all proofs are similar, we prove only (iv).

Assume that A is subregular. Let $a, b, c \in A$ and let α be the smallest congruence on A having $[a']\Theta(b, c)$ as a block for all $a' \in \langle a \rangle$. Then for all $a' \in \langle a \rangle$ we have $[a']\alpha = [a']\Theta(b, c)$, whence $\alpha = \Theta(b, c)$ by subregularity. Thus

$$\Theta(b, c) = \bigvee(\Theta(a', d) \mid a' \in \langle a \rangle \;\&\; d \in [a']\Theta(b, c)).$$

Since $\Theta(b, c)$ is compact, there exist $a_1, \ldots, a_n \in \langle a \rangle$ and $d_i \in [a_i]\Theta(b, c)$ with

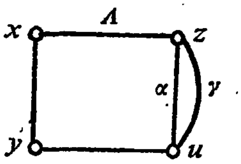
$$\Theta(b, c) = \bigvee_{i=1}^{n} \Theta(a_i, d_i).$$

Conversely, suppose that the principal-congruence condition holds. Let $B \leqq A$ and suppose that $\alpha, \beta \in \operatorname{Con} A$ satisfy $[b]\alpha \subseteq [b]\beta$ for all $b \in B$. Let $a \in B$ and $b, c \in A$ and let $a_i \in \langle a \rangle \subseteq B$ and $d_{ij} \in A$ be given by the principal-congruence condition. Now suppose that $b \equiv c(\alpha)$. Since $\Theta(a_i, d_i) \subseteq \Theta(b, c) \subseteq \alpha$, we have $d_i \in [a_i]\alpha \subseteq [a_i]\beta$ for all i and hence $a_i \equiv d_i(\beta)$ for all i. Thus $\Theta(b, c) \subseteq \beta$, whence $b \equiv c(\beta)$. Consequently $\alpha \subseteq \beta$.

While the proof below of congruence modularity uses Gumm's Shifting Principle, it is closely related to the corresponding proof in Bulman-Fleming, Day and Taylor [1].

1.4. Theorem. *If every subalgebra of* A^2 *is subregular, then* Con A *is modular.*

Proof. By Lemma 3.2 of Gumm [9] it suffices to prove that if $\alpha, \gamma \in \mathrm{Con}\, A$ and $\Lambda \leqq A^2$ is reflexive and symmetric with $\alpha \cap \Lambda \leqq \gamma \leqq \alpha$, then whenever we have



it follows that $x\gamma y$. Let α, γ and Λ be as stated and assume that the relations indicated in the diagram hold.

Consider $\alpha \times \gamma$ and $\gamma \times \gamma$ as congruences on the subregular algebra Λ: note that $\gamma \times \gamma \leqq \alpha \times \gamma$. Denote the diagonal of A^2 by Δ and consider $(a, a) \in \Delta$. Let $(b, c) \in \Lambda$ with $(b, c)\ \alpha \times \gamma\ (a, a)$. Then

$$
\begin{aligned}
b\,\alpha\,a \quad \& \quad c\,\gamma\,a \quad \& \quad b\,\Lambda\,c &\Rightarrow b\,\alpha\cap\Lambda\,c && \text{as} \quad \gamma \leqq \alpha \\
&\Rightarrow b\,\gamma\,c && \text{as} \quad \alpha\cap\Lambda \leqq \gamma \\
&\Rightarrow b\,\gamma\,a && \text{as} \quad c\,\gamma\,a \\
&\Rightarrow (b, c)\,\gamma\times\gamma\,(a, a).
\end{aligned}
$$

Thus $[(a, a)]_\Lambda \gamma \times \gamma = [(a, a)]_\Lambda \alpha \times \gamma$.

Hence $B := [\Delta]_\Lambda \gamma \times \gamma = [\Delta]_\Lambda \alpha \times \gamma$ and $(\gamma \times \gamma){\restriction}B = (\alpha \times \gamma){\restriction}B$. Consequently $\gamma \times \gamma = \alpha \times \gamma$ on Λ as Λ is subregular. Since $(x, z), (y, u) \in \Lambda$ with $(x, z)\ \alpha \times \gamma\ (y, u)$ we have $x\ \gamma\ y$, as required. (Note that the symmetry of Λ was not required.)

Similarly it can be proved that if $\mathrm{S}(A^2) \models \mathrm{R}(o_1, \ldots, o_n)$ then Con A is modular. It follows trivially from Theorem 1.4 that if $\mathrm{S}(A^2) \models \mathrm{SR}_n$ for some n, then Con A is modular; it seems highly unlikely that a similar conclusion can be made about R_n since the elements with respect to which A is regular cannot be forced into the diagonal.

2. Global relationships. In [2], Csákány characterized regularity for varieties via a quasi-identity: a variety $\mathcal{V}$ is regular if and only if there are ternary terms $p_1, \ldots, p_n$ such that

$$\mathcal{V} \models x = y \leftrightarrow \mathop{\&}_{i=1}^{n} p_i(xyz) = z.$$

Much earlier, Thurston [14] showed that $\mathcal{V}$ is regular if and only if for all $A \in \mathcal{V}$, all $\alpha \in \mathrm{Con}\, A$ and all $a \in A$ we have that $|[a]\alpha| = 1$ implies $\alpha = 0$. We now give the corresponding characterizations for our more general forms of regularity. Along the way we shall see that at the global level the various regularities come closer together.

The following translation principle allows us to convert a Hashimoto-type principal-congruence property directly into a quasi-identity.

2.1. Theorem. *Let $\mathcal{V}$ be a variety and let f_i, g_i, r and s be n-ary terms. Then the following are equivalent:*

(i) $\mathcal{V} \models \big(\mathop{\&}_{i=1}^{m} f_i(\vec{x}) = g_i(\vec{x})\big) \to r(\vec{x}) = s(\vec{x})$;

(ii) *for all* $A \in \mathcal{V}$ *and all* $\vec{a} \in A^n$

$$\bigvee_{i=1}^{m} \Theta\big(f_i(\vec{a}), g_i(\vec{a})\big) \supseteq \Theta\big(r(\vec{a}), s(\vec{a})\big);$$

(iii) *the elements* $f_i(\vec{x}), g_i(\vec{x}), r(\vec{x})$ *and* $s(\vec{x})$ *of* $F\mathcal{V}(n)$ *satisfy*

$$\bigvee_{i=1}^{m} \Theta\big(f_i(\vec{x}), g_i(\vec{x})\big) \supseteq \Theta\big(r(\vec{x}), s(\vec{x})\big).$$

Proof. (i)⇒(ii). Let $A \in \mathcal{V}$ and $\vec{a} \in A^n$ and define α to be $\bigvee_{i=1}^{m} \Theta\big(f_i(\vec{a}), g_i(\vec{a})\big)$. In A/α we have $f_i(\vec{b}) = g_i(\vec{b})$ for all i where $b_j := [a_j]\alpha$. Thus, by (i), $r(\vec{b}) = s(\vec{b})$, whence $r(\vec{a})\ \alpha\ s(\vec{a})$ as required.

(ii)⇒(iii) is trivial.

(iii)⇒(i). Let $A \in \mathcal{V}$ and $\vec{a} \in A^n$ with $f_i(\vec{a}) = g_i(\vec{a})$ for all i. Let $\varphi\colon F\mathcal{V}(n) \to A$ be a homomorphism with $x_j\varphi = a_j$. Then $\bigvee_{i=1}^{m} \Theta\big(f_i(\vec{x}), g_i(\vec{x})\big) \subseteq \ker\varphi$ and so $r(\vec{x}) \ker\varphi\ s(\vec{x})$ by (iii). Thus $r(\vec{a}) = s(\vec{a})$, as required.

In the following result we require the observation that if A is regular with respect to $a_1, \dots, a_n \in A$ and $\varphi\colon A \to B$ is a surjective homomorphism with $b_i = a_i\varphi$, then B is regular with respect to $b_1, \dots, b_n$.

2.2. Theorem. *The following are equivalent for any variety* $\mathcal{V}$:

(i) $\mathcal{V} \models \mathsf{SR}_n$;

(ii) $\mathcal{V} \models \mathsf{R}_n$;

(iii) *there exist unary terms* $u_1, \dots, u_n$ *such that for all* $A \in \mathcal{V}$ *and all* $a \in A$, A *is regular with respect to* $u_1(a), \dots, u_n(a)$;

(iv) *there exist unary terms* $u_1, \dots, u_n$ *such that for all* $A \in \mathcal{V}$, *all* $a \in A$ *and all* $\alpha \in \operatorname{Con} A$ *if* $|[u_1(a)]\alpha| = \dots = |[u_n(a)]\alpha| = 1$, *then* $\alpha = 0$;

(v) *there exist unary terms* $u_1, \dots, u_n$ *and ternary terms* $p_{11}, \dots, p_{nm}$ *such that*

$$\mathcal{V} \models \big(\mathop{\&}_{i=1}^{n} \mathop{\&}_{j=1}^{m} p_{ij}(xyz) = u_i(z)\big) \leftrightarrow x = y;$$

(vi) *there exist unary terms* $u_1, \ldots, u_n$ *and ternary terms* $p_1, \ldots, p_m$ *and a selection function* $j \mapsto i_j$ *such that*

$$\mathscr{V} \models \Big(\mathop{\&}_{j=1}^{m} p_j(xyz) = u_{i_j}(z)\Big) \leftrightarrow x = y.$$

Proof. That (i) implies (ii) is trivial. Assume that $\mathscr{V} \models \mathsf{R}_n$. Then there exist $v_1, \ldots, v_n \in F\mathscr{V}(\mathrm{N})$ such that $F\mathscr{V}(\mathrm{N})$ is regular with respect to $v_1, \ldots, v_n$. Assume that $v_1, \ldots, v_n$ depend only upon $x_1, \ldots, x_k$; then we can find an onto homomorphism ψ: $F\mathscr{V}(\mathrm{N}) \to F\mathscr{V}(\{x, y, z\})$ with $x_i\varphi = z$ for $i = 1, \ldots, k$. Thus the image u_i of v_i under φ depends only upon z and $F\mathscr{V}(3)$ is regular with respect to $u_1, \ldots, u_n$.

Suppose that $A \in \mathscr{V}$, $a \in A$ and $\alpha, \beta \in \mathrm{Con}\, A$ with $[u_i(a)]\alpha \subseteq [u_i(a)]\beta$ for all i. Let $s, t \in A$ with $s\ \alpha\ t$ and define φ: $F\mathscr{V}(\{x, y, z\}) \to A$ by $x\varphi = s$, $y\varphi = t$, $z\varphi = a$. Then $x\ \bar{\alpha}\ y$ where $\bar{\alpha}$ denotes the inverse image of α under φ. Now $v\ \bar{\alpha}\ u_i(z)$ implies $v\varphi\ \alpha\ u_i(a)$; hence $v\varphi\ \beta\ u_i(a)$ and so $v\ \bar{\beta}\ u_i(z)$. Thus $[u_i(z)]\bar{\alpha} \subseteq [u_i(z)]\bar{\beta}$ for all i and consequently $\bar{\alpha} \subseteq \bar{\beta}$ since $F\mathscr{V}(\{x, y, z\})$ is regular with respect to $u_1(z), \ldots, u_n(z)$. Hence

$$s\ \alpha\ t \Rightarrow x\ \bar{\alpha}\ y \Rightarrow x\ \bar{\beta}\ y \Rightarrow s\ \beta\ t$$

and thus $\alpha \subseteq \beta$. Hence (ii) implies (iii).

That (v) follows from (iii) is a direct consequence of the principal-congruence characterization of regularity with respect to $a_1, \ldots, a_n$ given in Lemma 1.3 (i) and the translation to quasi-identities given in Theorem 2.1: take $A = F\mathscr{V}(\{x, y, z\})$, $a_i = u_i(z)$, $b = x$, $c = y$ and $d_{ij} = p_{ij}(xyz)$. The equivalence of (v) and (vi) is clear.

The combination of Theorem 2.1 and Lemma 1.3 (v) shows that (v) implies (i). It remains to prove that (iv) implies (iii).

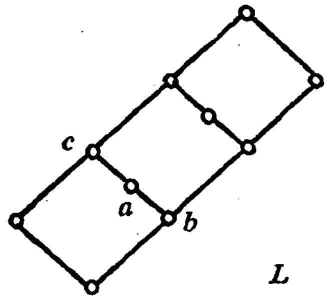
Suppose that $[u_i(a)]\alpha \subseteq [u_i(a)]\beta$ for $i = 1, \ldots, n$; then $[u_i(a)]\alpha = [u_i(a)]\alpha \wedge \beta$. Consider the congruence $\alpha/(\alpha \wedge \beta)$ on $A/(\alpha \wedge \beta)$. The block of $[u_i(a)]\alpha \wedge \beta$ in $\alpha/(\alpha \wedge \beta)$ is a singleton for all i and hence, by (iv), we have $\alpha/(\alpha \wedge \beta) = 0$ in $\mathrm{Con}\, A/(\alpha \wedge \beta)$. Thus $\alpha = \alpha \wedge \beta$ and so $\alpha \subseteq \beta$, as required.

The choice between (v) and (vi) is a matter of taste: in (v) the emphasis is on the unary terms while the emphasis in (vi) is on the ternary terms. The equivalence of SR_1 and R_1 was observed by GRÄTZER [8]. It is tempting to replace (iv) by

(iv)′ $(\forall A \in \mathscr{V})(\exists a_1, \ldots, a_n \in A)(\forall \alpha \in \mathrm{Con}\, A)(|[a_1]\alpha| = \ldots = |[a_n]\alpha| = 1) \Rightarrow \alpha = 0.$

An algebra A with this property might be called *regular at* 0 *with respect to* $a_1, \ldots, a_n$. But this property is not preserved by homomorphisms and so the proof method used above is not applicable. The lattice L drawn below is regular at 0 with respect to a as it is subdirectly irreducible and both ab and ac are critical edges (that is,

$\Theta(a, b)=\Theta(a, c)$ is the monolith of L). Since $L/\Theta(a, b)$ is a four-element chain it is not regular at 0 with respect to any one of its elements.



L

The proof of our next result is now easy and is omitted.

2.3. Theorem. *The following are equivalent for any variety* $\mathscr{V}$:

(i) $\mathscr{V} \models \mathsf{SR}$;

(ii) $(\exists n\in \mathrm{N})\mathscr{V} \models \mathsf{SR}_n$; (ii)′ $(\forall A\in\mathscr{V})(\exists n\in\mathrm{N}) A\models \mathsf{SR}_n$;

(iii) $(\exists n\in \mathrm{N})\mathscr{V} \models \mathsf{R}_n$; (iii)′ $(\forall A\in\mathscr{V})(\exists n\in\mathrm{N}) A\models \mathsf{R}_n$;

(iv) *for all* $A\in\mathscr{V}$, *all* $B\leqq A$ *and all* $\alpha\in\operatorname{Con} A$ *if* $|[b]\alpha|=1$ *for all* $b\in B$, *then* $\alpha=0$;

(v) *there exist* $n\in\mathrm{N}$, *unary terms* $u_1, \ldots, u_n$ *and ternary terms* $p_1, \ldots, p_n$ *such that*

$$\mathscr{V} \models \left(\mathop{\&}_{i=1}^{n} p_i(xyz) = u_i(z)\right) \leftrightarrow x = y.$$

Call a variety $\mathscr{V}$ *locally regular with respect to unary terms* $u_1, \ldots, u_n$ (and write $\mathscr{V} \models \mathsf{LR}(u_1, \ldots, u_n)$) if

$$(\forall A\in\mathscr{V})(\forall a\in A)\left[\left(\mathop{\&}_{i=1}^{n} [u_i(a)]\alpha = [u_i(a)]\beta\right) \Rightarrow [a]\alpha = [a]\beta\right].$$

This concept was introduced, under a different name, in the important but unpublished paper Hagemann [10] where a characterization via identities was obtained; no quasi-identity characterization was given. It is clear from Theorem 2.2 that, at the varietal level, we have $\mathsf{R}_n \to \mathsf{LR}_n$. The proof of the following result should by now be an easy exercise.

2.4. Theorem. *The following are equivalent for any variety* $\mathscr{V}$ *and unary terms* $u_1, \ldots, u_n$:

(i) $\mathscr{V} \models \mathsf{LR}(u_1, \ldots, u_n)$;

(ii) *for all* $A\in\mathscr{V}$, *all* $a\in A$ *and all* $\alpha, \beta\in\operatorname{Con} A$

$$\mathop{\&}_{i=1}^{n} \left([u_i(a)]\alpha \subseteqq [u_i(a)]\beta\right) \Rightarrow [a]\alpha \subseteqq [a]\beta;$$

(iii) *for all* $A \in \mathcal{V}$ *and all* $a, b \in A$ *there exist* $d_{i1}, \ldots, d_{im} \in A$ *such that*

$$\Theta(a, b) = \bigvee_{i=1}^{n} \bigvee_{j=1}^{m} \Theta(u_i(a), d_{ij});$$

(iv) *for all* $A \in \mathcal{V}$, *all* $a \in A$ *and all* $\alpha \in \operatorname{Con} A$ *if* $|[u_1(a)]\alpha| = \ldots = |[u_n(a)]\alpha| = 1$, *then* $|[a]\alpha| = 1$;

(v) *there exist binary terms* $p_{11}, \ldots, p_{nm}$ *such that*

$$\mathcal{V} \models (\mathop{\&}_{i=1}^{n} \mathop{\&}_{j=1}^{m} p_{ij}(xy) = u_i(x)) \leftrightarrow x = y;$$

(vi) *there exist binary terms* $p_1, \ldots, p_k$ *and a selection function* $j \mapsto i_j$ *such that*

$$\mathcal{V} \models (\mathop{\&}_{j=1}^{k} p_j(xy) = u_{i_j}(x)) \leftrightarrow x = y.$$

Note that $\mathsf{R}(o_1, \ldots, o_n)$ for constant terms $o_1, \ldots, o_n$ implies LR_n and if the terms in the definition of LR_n can be chosen to be constants then we obtain the reverse implication. Thus Theorem 2.4 yields the quasi-identity characterization of $\mathsf{R}(o_1, \ldots, o_n)$.

2.5. Corollary. *The following are equivalent for any variety* $\mathcal{V}$ *and constant terms* $o_1, \ldots, o_n$:

(i) $\mathcal{V} \models \mathsf{R}(o_1, \ldots, o_n)$;

(ii) *there exist binary terms* $p_{11}, \ldots, p_{nm}$ *such that*

$$\mathcal{V} \models (\mathop{\&}_{i=1}^{n} \mathop{\&}_{j=1}^{m} p_{ij}(xy) = o_i) \leftrightarrow x = y;$$

(iii) *there exist binary terms* $p_1, \ldots, p_k$ *and a selection function* $j \mapsto i_j$ *such that*

$$\mathcal{V} \models (\mathop{\&}_{j=1}^{k} p_j(xy) = o_{i_j}) \leftrightarrow x = y.$$

3. Quasi-identities, congruence modularity and permutability. In this section we give the general translation from quasi-identities to identities and investigate the relationship between quasi-identities, congruence modularity and n-permutability.

Lemmas 3.1 and 3.2 are simply restatements of Mal'cev's description of principal congruences. If $Z \subseteq A^2$, then $\Theta(Z)$ denotes the smallest congruence containing Z.

3.1. Lemma. *Let* $Z \subseteq A^2$ *and let* $(c, d) \in A^2$. *Then* $(c, d) \in \Theta(Z)$ *if and only if for some* $k, l \in \mathbf{N}$ *there exist* $(l+2)$-*ary terms* $w_1, \ldots, w_k$, *there exists* $\bar{e} \in A^l$ *and* *there are pairs* (a_i, b_i) *such that*

$$c = w_1(a_1, b_1, \vec{e})$$

$$w_1(b_1, a_1, \vec{e}) = w_2(a_1, b_1, \vec{e})$$

$$\vdots$$

$$w_k(b_k, a_k, \vec{e}) = d,$$

and $(a_i, b_i) \in Z$ *for all* i.

3.2. Lemma. *Let* $Z \subseteq A^2$ *and let* $(c, d) \in A^2$. *Then* $(c, d) \in \Theta(Z)$ *if and only if for some* $k, l \in \mathbf{N}$ *there exist* $(l+1)$*-ary terms* $w_1, \ldots, w_k$, *there exists* $\vec{e} \in A^l$ *and there are pairs* (a_i, b_i) *such that*

$$c = w_1(a_1, \vec{e})$$

$$w_1(b_1, \vec{e}) = w_2(a_2, \vec{e})$$

$$\vdots$$

$$w_k(b_k, \vec{e}) = d,$$

and $(a_i, b_i) \in Z$ *or* $(b_i, a_i) \in Z$ *for all* i.

Recall that A is called *k-permutable* if for all $\alpha, \beta \in \mathrm{Con}\, A$ we have $\alpha \vee \beta = \alpha \circ \beta \circ \alpha \ldots$ (with k factors).

Clearly the last line of Lemma 3.2 is needed to guarantee symmetry. Hagemann [10] showed that if $\mathscr{V}$ is a k-permutable variety then for all $A \in \mathscr{V}$, if R is a reflexive subalgebra of A^2, then $R \circ \ldots \circ R$ (with $k-1$ factors) is a congruence. Using this we can simplify Lemma 3.2. The result was rediscovered by Lakser [13] and Duda [5].

3.3. Lemma. *Assume that* A *belongs to a* $(k+1)$*-permutable variety. Let* $Z \subseteq A^2$ *and let* $(c, d) \in A^2$. *Then* $(c, d) \in \Theta(Z)$ *if and only if for some* $l \in \mathbf{N}$ *there exist* $(l+1)$*-ary terms* $w_1, \ldots, w_k$, *there exists* $\vec{e} \in A^l$ *and there are pairs* (a_i, b_i) *such that*

$$c = w_1(a_1, \vec{e})$$

$$w_1(b_1, \vec{e}) = w_2(a_2, \vec{e})$$

$$\vdots$$

$$w_k(b_k, \vec{e}) = d,$$

and $(a_i, b_i) \in Z$ *for all* i.

The translation from quasi-identities to identities is obtained by combining one of these lemmas with the principal-congruence translation given in the previous section.

3.4. Theorem. *Let $\mathscr{V}$ be a variety and let f_i, g_i, r and s be n-ary terms. Then the following are equivalent:*

(i) $\mathscr{V} \models \left(\underset{i=1}{\overset{m}{\&}} f_i(\vec{x}) = g_i(\vec{x})\right) \to r(\vec{x}) = s(\vec{x})$;

(ii) *for some* $k \in \mathrm{N}$ *there exist* $(n+2)$*-ary terms* $t_1, \ldots, t_k$ *and pairs* $(u_j, v_j) \in \{(f_i, g_i) | i=1, \ldots, m\}$ *such that* $\mathscr{V}$ *satisfies the identities*

$$r(\vec{x}) = t_1(u_1(\vec{x}), v_1(\vec{x}), \vec{x})$$
$$t_1(v_1(\vec{x}), u_1(\vec{x}), \vec{x}) = t_2(u_2(\vec{x}), v_2(\vec{x}), \vec{x})$$
$$\vdots$$
$$t_k(v_1(\vec{x}), u_2(\vec{x}), \vec{x}) = s(\vec{x});$$

(iii) *for some* $k \in \mathrm{N}$ *there exist* $(n+1)$*-ary terms* $t_1, \ldots, t_k$ *and pairs* $(u_j, v_j) \in \{(f_i, g_i), (g_i, f_i) \mid i=1, \ldots, m\}$ *such that* $\mathscr{V}$ *satisfies*

$$r(\vec{x}) = t_1(u_1(\vec{x}), \vec{x})$$
$$t_1(v_1(\vec{x}), \vec{x}) = t_2(u_2(\vec{x}), \vec{x})$$
$$\vdots$$
$$t_k(v_k(\vec{x}), \vec{x}) = s(\vec{x}).$$

Proof. Assume that (i) holds. Then by Theorem 2.1 we have $r(\vec{x}) \equiv s(\vec{x})(\Theta(Z))$ on $F\mathscr{V}(n)$ where $Z = \{(f_i, g_i) \mid i=1, \ldots, m\}$. Thus by Lemma 3.1, for some $k, l \in \mathrm{N}$ there exist $(l+2)$-ary terms $w_1, \ldots, w_k$ and n-ary terms $h_1, \ldots, h_l$ and pairs $(u_j, v_j) \in Z$ such that (in $F\mathscr{V}(n)$)

$$r(\vec{x}) = w_1(u_1(\vec{x}), v_1(\vec{x}), h_1(\vec{x}), \ldots, h_l(\vec{x}))$$
$$w_1(v_1(\vec{x}), u_1(\vec{x}), h_1(\vec{x}), \ldots, h_l(\vec{x})) = w_2(u_2(\vec{x}), v_2(\vec{x}), h_1(\vec{x}), \ldots, h_l(\vec{x}))$$
$$\vdots$$
$$w_k(v_k(\vec{x}), u_k(\vec{x}), h_1(\vec{x}), \ldots, h_l(\vec{x})) = s(\vec{x}).$$

Thus (ii) holds: define $t_j(y, z, \vec{x}) = w_j(y, z, h_1(\vec{x}), \ldots, h_l(\vec{x}))$. That (ii) implies (i) is trivial. In the same way, Lemma 3.2 yields the equivalence of (i) and (iii).

In just the same way, Theorem 2.1 and Lemma 3.3 combine to yield a simpler Mal'cev condition in the $(k+1)$-permutable case.

3.5. Theorem. *Let $\mathscr{V}$ be a $(k+1)$-permutable variety and let f_i, g_i, r and s be n-ary terms. Then the following are equivalent:*

(i) $\mathscr{V} \models \left(\underset{i=1}{\overset{m}{\&}} f_i(\vec{x}) = g_i(\vec{x})\right) \to r(\vec{x}) = s(\vec{x})$;

(ii) *there exist* $(n+1)$*-ary terms* $t_1, \ldots, t_k$ *and pairs* $(u_j, v_j) \in \{(f_i, g_i) \mid i=1, \ldots, m\}$

such that $\mathscr{V}$ *satisfies the identities*

$$\begin{aligned} r(\vec{x}) &= t_1(u_1(\vec{x}), \vec{x}) \\ t_1(v_1(\vec{x}), \vec{x}) &= t_2(u_2(\vec{x}), \vec{x}) \\ &\vdots \\ t_k(v_k(\vec{x}), \vec{x}) &= s(\vec{x}). \end{aligned}$$

If a variety $\mathscr{V}$ is k-permutable we shall write $\mathscr{V} \models \mathsf{P}_k$ and if $\mathscr{V}$ is k-permutable for some $k \in \mathbf{N}$ then we write $\mathscr{V} \models \mathsf{P}_*$. Whenever every algebra in $\mathscr{V}$ has a modular congruence lattice we write $\mathscr{V} \models \mathsf{CM}$. We require the identities for k-permutability (Hagemann and Mitschke [11]) and for congruence modularity Day [4].

3.6. Lemma. *Let* $\mathscr{V}$ *be a variety.*

(a) *Let* $k \geqq 2$. *Then* $\mathscr{V} \models \mathsf{P}_k$ *if and only if there are 3-ary terms* $p_1, \ldots, p_{k-1}$ *such that* $\mathscr{V}$ *satisfies*

$$\begin{aligned} x &= p_1(xzz), \\ p_i(xxz) &= p_{i+1}(xzz) \quad \textit{for all } i, \\ p_{k-1}(xxz) &= z. \end{aligned}$$

(b) $\mathscr{V} \models \mathsf{CM}$ *if and only if for some* $n \geqq 2$ *there exist 4-ary terms* $m_0, \ldots, m_n$ *such that* $\mathscr{V}$ *satisfies*

$$\begin{aligned} m_0(xyzw) &= x, & \\ m_i(xyyx) &= x & \textit{for all } i, \\ m_i(xxww) &= m_{i+1}(xxww) & \textit{for even } i, \\ m_i(xyyw) &= m_{i+1}(xyyw) & \textit{for odd } i, \\ m_n(xyzw) &= w. & \end{aligned}$$

(c) $\mathscr{V} \models \mathsf{CM}$ *if and only if for some* $n \geqq 2$ *there exist 4-ary terms* $m_0', \ldots, m_n'$ *such that* $\mathscr{V}$ *satisfies*

$$\begin{aligned} m_0'(xyzw) &= x, & \\ m_i'(xyyx) &= x & \textit{for all } i, \\ m_i'(xyyw) &= m_{i+1}'(xyyw) & \textit{for even } i, \\ m_i'(xxww) &= m_{i+1}'(xxww) & \textit{for odd } i, \\ m_n'(xyzw) &= w. & \end{aligned}$$

When the condition given in (b) above holds we shall write $\mathscr{V} \models \mathsf{CM}_n$. Similarly for the condition in (c) we write $\mathscr{V} \models \mathsf{CM}_n'$. For $n=2$ the m_i and m_i' are interdefinable but do not seem to be for $n \geqq 3$. Clearly $\mathsf{CM}_n \Rightarrow \mathsf{CM}_{n+1}' \Rightarrow \mathsf{CM}_{n+2}$

and hence $\mathsf{V}\,\mathsf{CM}_n = \mathsf{V}\,\mathsf{CM}'_n = \mathsf{CM}$. We shall refer to the terms m_i and the terms m'_i as the *Day terms*.

3.7. Lemma. *Let $\mathscr{V}$ be a variety. If $\mathscr{V} \models \mathsf{CM}_n$ and the Day terms satisfy $m_i(xxxz) = m_i(xzzz)$ for all (even or odd) i, then $\mathscr{V} \models \mathsf{P}_n$. Similarly if $\mathscr{V} \models \mathsf{CM}'_n$ and the Day terms satisfy $m'_i(xxxz) = m'_i(xzzz)$ for all (even or odd) i, then $\mathscr{V} \models \mathsf{P}_n$.*

Proof. Assume that $\mathscr{V} \models \mathsf{CM}_n$ with $m_i(xxxz) = m_i(xzzz)$ for all i. Define 3-ary terms $p_1, \ldots, p_{n-1}$ by

$$p_i(xyz) = \begin{cases} m_i(xxyz) & \text{for odd } i, \\ m_i(xyzz) & \text{for even } i. \end{cases}$$

Then by Lemma 3.6 (b) and our extra assumption, we have

$$p_1(xzz) = m_1(xxzz) = m_0(xxzz) = x,$$

$$(i \text{ odd}) \quad p_i(xxz) = m_i(xxxz) = m_i(xzzz) = m_{i+1}(xzzz) = p_{i+1}(xzz),$$

$$(i \text{ even}) \quad p_i(xxz) = m_i(xxzz) = m_{i+1}(xxzz) = p_{i+1}(xzz),$$

$$p_{n-1}(xxz) = p_n(xzz) = m_n(\ldots z) = z.$$

Thus, by Lemma 3.6 (a), we have $\mathscr{V} \models \mathsf{P}_n$. The proof for $\mathsf{CM}'_n \to \mathsf{P}_n$ is similar.

3.8. Lemma. *On any variety we have:*

(i) $\mathsf{CM}_2 \leftrightarrow \mathsf{CM}'_2 \leftrightarrow \mathsf{P}_2$;

(ii) $\mathsf{CM}'_3 \leftrightarrow \mathsf{P}_3$.

Proof. (i) Let m_1 be the nontrivial term for CM_2. Then $m'_1(xyzw) := m_1(wzyx)$ is the corresponding term for CM'_2. Thus $\mathsf{CM}_2 \to \mathsf{CM}'_2$ and similarly $\mathsf{CM}'_2 \to \mathsf{CM}_2$. The term for P_2 is given by $p_1(xyz) := m_1(xxyz)$. Thus $\mathsf{CM}_2 \to \mathsf{P}_2$ and the converse holds by the previous lemma since $m_1(xxxz) = m_2(xxxz) = z = m_2(xzzz) = m_1(xzzz)$.

(ii) It is easily seen that the Day terms for CM'_3 satisfy $m'_i(xxxz) = m'_i(xzzz)$ and hence $\mathsf{CM}'_3 \to \mathsf{P}_3$ by the previous lemma. If p_1 and p_2 are the terms for P_3 then terms m'_1 and m'_2 for CM'_2 may be defined by $m'_1(xyzw) := p_1(xyz)$ and $m'_2(xyzw) := {}$ $:= p_2(yzw)$; the identities for CM'_2 are easily checked. Thus $\mathsf{P}_3 \to \mathsf{CM}'_3$.

Hagemann [10] observed that for varieties we have $\mathsf{R} \to \mathsf{CM}$ and $\mathsf{R} \to \mathsf{P}_*$. Since regularity is given by a quasi-identity, it is natural to ask which quasi-identities yield CM and P_*.

3.9. Theorem. *Let $\mathscr{V}$ be a variety, let f_i, g_i be $(n+2)$-ary terms $(n \geqq 0)$ and let h_i be unary terms such that $\mathscr{V}$ satisfies*

$$\Big(\mathop{\&}_{i=1}^{m} f_i(xy\bar{z}) = g_i(xy\bar{z})\Big) \to x = y$$

and

$$f_i(xx\underline{z}) = g_i(xx\underline{z}) = h_i(z),$$

where $\underline{z}=(z...z)$. *Then* $\mathscr{V} \models \mathsf{CM} \,\&\, \mathsf{P}_*$.

Proof. Assume that $\mathscr{V}$ satisfies the quasi-identity and identities above, and let $t_1, \ldots, t_k$ be the $(n+4)$-ary terms given by Theorem 3.4. Thus there are pairs $(u_j, v_j) \in \{(f_i, g_i) \mid i=1, \ldots, m\}$ such that $\mathscr{V}$ satisfies

$$x = t_1(u_1(xy\vec{z}), v_1(xy\vec{z}), xy\vec{z})$$

$$t_1(v_1(xy\vec{z}), u_1(xy\vec{z}), xy\vec{z}) = t_2(u_2(xy\vec{z}), v_2(xy\vec{z}), xy\vec{z})$$

$$\vdots$$

$$t_k(v_k(xy\vec{z}), u_k(xy\vec{z}), xy\vec{z}) = y,$$

and there exist unary terms $w_j \in \{h_1, \ldots, h_m\}$ such that

$$u_j(xx\underline{z}) = v_j(xx\underline{z}) = w_j(z).$$

We shall prove that $\mathscr{V} \models \mathsf{CM}_{2k+1} \,\&\, \mathsf{P}_{2k+1}$. Define the Day terms as follows:

$$m_0(xyzw) = x,$$

$$m_{2j-1}(xyzw) = t_j(u_j(yz\underline{w}), v_j(yz\underline{w}), xw\underline{w}),$$

$$m_{2j}(xyzw) = t_j(v_j(yz\underline{w}), u_j(yz\underline{w}), xw\underline{w}),$$

$$m_{2k+1}(xyzw) = w.$$

Rather than introduce w_j into the calculations we shall repeatedly use the fact that $u_j(xx\underline{z})$ and $v_j(xx\underline{z})$ are equal and independent of x. For $0<j<k$ we have

$$m_{2j-1}(xyyx) = t_j(u_j(yy\underline{x}), v_j(yy\underline{x}), xx\underline{x}) =$$

$$= t_j(v_j(yy\underline{x}), u_j(yy\underline{x}), xx\underline{x}) = m_{2j}(xyyx) =$$

$$= t_j(v_j(xx\underline{x}), u_j(xx\underline{x}), xx\underline{x}) = t_{j+1}(u_{j+1}(xx\underline{x}), v_{j+1}(xx\underline{x}, xx\underline{x}) =$$

$$= t_{j+1}(u_{j+1}(yy\underline{x}), v_{j+1}(yy\underline{x}), xx\underline{x}) = m_{2j+1}(xyyx).$$

A similar calculation shows that $m_1(xyyx)=x$ and it follows by induction that $m_i(xyyx)=x$ for all i. Now

$$m_0(xxww) = x = t_1(u_1(xw\underline{w}), v_1(xw\underline{w}), xw\underline{w}) = m_1(xxww)$$

and similarly

$$m_{2k}(xxww) = t_k(v_k(xw\underline{w}), u_k(xw\underline{w}), xw\underline{w}) = w = m_{2k+1}(xxww),$$

and for $0<j<k$ we find

$$m_{2j}(xxww) = t_j(v_j(xw\underline{w}), u_j(xw\underline{w}), xw\underline{w}) =$$

$$= t_{j+1}(u_{j+1}(xw\underline{w}), v_{j+1}(xw\underline{w}), xw\underline{w}) = m_{2j+1}(xxww).$$

Hence $m_i(xxww)=m_{i+1}(xxww)$ for i even. Finally, for $1\leq j\leq k$ we have

$$m_{2j-1}(xyyw) = t_j(u_j(yy\underline{w}), v_j(yy\underline{w}), xw\underline{w}) =$$
$$= t_j(v_j(yy\underline{w}), u_j(yy\underline{w}), xw\underline{w}) = m_{2j}(xyyw),$$

and thus $m_i(xyyw)=m_{i+1}(xyyw)$ for i odd. Consequently $\mathcal{V}\models \mathsf{CM}_{2k+1}$ by Lemma 3.6 (b).

By Lemma 3.7, to show that $\mathcal{V}\models \mathsf{P}_{2k+1}$ it suffices to show that the Day terms defined above satisfy $m_i(xxxz)=m_i(xzzz)$ for odd i (and hence for all i). But for $1\leq j\leq k$ we find

$$m_{2j-1}(xxxz) = t_j(u_j(xx\underline{z}), v_j(xx\underline{z}), xz\underline{z}) =$$
$$= t_j(u_j(zz\underline{z}), v_j(zz\underline{z}), xz\underline{z}) = m_{2j-1}(xzzz),$$

as required.

These considerations lead us to ask for compact collections of identities characterizing $\mathsf{CM}\,\&\,\mathsf{P}_*$ and $\mathsf{CM}\,\&\,\mathsf{P}_k$. Note that $\mathsf{CM}\,\&\,\mathsf{P}_k$ is equivalent to $\mathsf{CM}_k\,\&\,\mathsf{P}_k$. Our Lemma 3.7 gives a useful set of identities which imply $\mathsf{CM}_k\,\&\,\mathsf{P}_k$ while Lemma 3.8 shows that there is noting to do for $k=2, 3$.

4. Applications to congruence regularity. It is a simple exercise to apply the results of Section 3 to the various forms of regularity (and we leave all of the details to the reader). For example, we obtain at once that, at the varietal level,

$$(\mathsf{R}(o_1, \dots, o_n) \text{ or } \mathsf{R}_n \text{ or } \mathsf{SR}) \rightarrow \mathsf{CM}\,\&\,\mathsf{P}_*.$$

Since every variety satisfies $\mathsf{LR}(x)$, Theorem 2.4 shows that in Theorem 3.9 we cannot drop the additional assumption that $f_i(xx\underline{z})$ and $g_i(xx\underline{z})$ are independent of x.

Combining Theorems 2.2 and 2.3 with Theorem 3.5 gives the identities which characterize R_n and SR.

4.1. Theorem. *Let $\mathcal{V}$ be a variety. Then $\mathcal{V}\models \mathsf{R}_n$ if and only if there exist unary terms $u_1, \dots, u_n$, and for some $k\in\mathbf{N}$ there are 4-ary terms $t_1, \dots, t_k$ and 3-ary terms $p_1, \dots, p_k$ and there is a selection function $j\mapsto i_j$ such that $\mathcal{V}$ satisfies*

$$\begin{aligned} x &= t_1(p_1(xyz), xyz)\\ t_1(u_{i_1}(z), xyz) &= t_2(p_2(xyz), xyz)\\ &\;\vdots\\ t_k(u_{i_k}(z), xyz) &= y, \end{aligned}$$

and $p_j(xxz)=u_{i_j}(z)$ for all j.

4.2. Theorem. *Let $\mathcal{V}$ be a variety. Then $\mathcal{V}\models \mathsf{SR}$ if and only if for some $n\in\mathbf{N}$ there exist unary terms $u_1, \dots, u_n$, 4-ary terms $t_1, \dots, t_n$ and 3-ary terms*

$p_1, \ldots, p_n$ *such that* $\mathcal{V}$ *satisfies*

$$\begin{aligned} x &= t_1(p_1(xyz), xyz) \\ t_1(u_1(z), xyz) &= t_2(p_2(xyz), xyz) \\ &\vdots \\ t_n(u_n(z), xyz) &= y, \end{aligned}$$

and $p_j(xxz) = u_j(z)$ *for all* j.

This characterization of subregularity and the quasi-identity characterization from Theorem 2.3 were obtained independently by DUDA [6, 7].

If we combine Theorems 2.4 and 3.4 to give identities for $\mathsf{LR}(u_1, \ldots, u_n)$ we do not immediately obtain the identities given by HAGEMANN [10]. Theorem 3.4 and the following lemma, whose proof we leave to the reader, provide the translation from our identities to his.

4.3. Lemma. *Let* $n \geqq 2$ *and* $l \geqq 0$. *The following are equivalent for a variety* $\mathcal{V}$:

(i) *there exist* $(n+l)$*-ary terms* $p_1, \ldots, p_s$ *and* $q_1, \ldots, q_s$ *such that* $\mathcal{V}$ *satisfies*

$$\Big(\mathop{\&}_{i=1}^{s} p_i(x_1 \ldots x_n y_1 \ldots y_l) = q_i(x_1 \ldots x_n y_1 \ldots y_l)\Big) \leftrightarrow x_1 = \ldots = x_n;$$

(ii) *there exist* $(n+l)$*-ary terms* $u_1, \ldots, u_t$ *and* $(l+1)$*-ary terms* $v_1, \ldots, v_t$ *such that* $\mathcal{V}$ *satisfies*

$$\Big(\mathop{\&}_{j=1}^{t} u_j(x_1 \ldots x_n y_1 \ldots y_l) = v_j(x_1 y_1 \ldots y_l)\Big) \leftrightarrow x_1 = \ldots = x_n.$$

Moreover the translation between (i) *and* (ii) *can be achieved in such a way that on* $\mathcal{V}$ *we have*

$$\begin{aligned} &\{p_i(x \ldots x y_1 \ldots y_l) = q_i(x \ldots x y_1 \ldots y_l) \mid i = 1, \ldots, s\} = \\ &= \{u_j(x \ldots x y_1 \ldots y_l) = v_j(x y_1 \ldots y_l) \mid j = 1, \ldots, t\}. \end{aligned}$$

4.4. Theorem. *Let* $\mathcal{V}$ *be a variety. Then the following are equivalent:*

(i) $\mathcal{V} \models \mathsf{LR}(u_1, \ldots, u_n)$;

(ii) *for some* $k \in \mathbf{N}$ *there exist* 4*-ary terms* $t_1, \ldots, t_k$ *and binary terms* $p_1, \ldots, p_k$ *and a selection function* $j \mapsto i_j$ *such that* $\mathcal{V}$ *satisfies*

$$\begin{aligned} x &= t_1(p_1(xy), u_{i_1}(x), xy) \\ t_1(u_{i_1}(x), p_1(xy), xy) &= t_2(p_2(xy), u_{i_2}(x), xy) \\ &\vdots \\ t_k(u_{i_k}(x), p_k(xy), xy) &= y, \end{aligned}$$

and $p_j(xx) = u_{i_j}(x)$ *for all* j;

(iii) *for some* $k \in \mathbf{N}$ *there exist* 3*-ary terms* $t_1, \ldots, t_k$ *and binary terms* $p_1, \ldots, p_k$

and $q_1, \ldots, q_k$ *and a selection function* $j \mapsto i_j$ *such that* $\mathcal{V}$ *satisfies*

$$\begin{aligned} x &= t_1(p_1(xy), xy) \\ t_1(q_1(xy), xy) &= t_2(p_2(xy), xy) \\ &\vdots \\ t_k(q_k(xy), xy) &= y, \end{aligned}$$

and $p_j(xx) = q_j(xx) = u_{i_j}(x)$ *for all* j.

Condition (iii) of this theorem is precisely the characterization of $\mathsf{LR}(u_1, \ldots, u_n)$ given in HAGEMANN [10].

References

[1] S. BULMAN-FLEMING—A. DAY—W. TAYLOR, Regularity and modularity of congruences, *Algebra Universalis*, **4** (1974), 58—60.

[2] B. CSÁKÁNY, Characterizations of regular varieties, *Acta Sci. Math.*, **31** (1970), 187—189.

[3] B. A. DAVEY—L. G. KOVÁCS, Absolute subretracts and weak injectives in congruence modular varieties, *Trans. Amer. Math. Soc.*, **297** (1986), 181—186.

[4] A. DAY, A characterization of congruence modularity for congruence lattices of algebras, *Canad. Math. Bull.*, **12** (1969), 167—173.

[5] J. DUDA, On two congruence schemes applied to Mal'cev-type theorems, *Ann. Univ. Sci. Budapest., Sect. Math.*, **26** (1983), 39—45.

[6] J. DUDA, Regularity of algebras with applications to congruence class geometry, *Arch. Math. (Brno)*, **19** (1983), 199—208.

[7] J. DUDA, Mal'cev conditions for varieties of subregular algebras, *Acta Sci. Math.*, to appear.

[8] G. GRÄTZER, Two Mal'cev-type theorems in universal algebra, *J. Combin. Theory*, **8** (1970), 334—342.

[9] H. P. GUMM, Geometrical methods in congruence modular varieties, *Mem. Amer. Math. Soc.*, **45** (1983), no. 286.

[10] J. HAGEMANN, On regular and weakly regular congruences, Technische Hochschule Darmstadt, preprint no. 75, June 1973 (unpublished except as part of the author's Dr. rer. nat. dissertation in 1975).

[11] J. HAGEMANN—A. MITSCHKE, On n-permutable congruences, *Algebra Universalis*, **3** (1973), 8—12.

[12] J. HASHIMOTO, Congruence relations and congruence classes in lattices, *Osaka Math. J.*, **15** (1963), 71—86.

[13] H. LAKSER, Principal congruences in n-permutable varieties, *Algebra Universalis*, **14** (1982), 64—67.

[14] H. A. THURSTON, Derived operations and congruences, *Proc. London Math. Soc.* (3), **8** (1958), 127—134.

[15] J. TIMM, On regular algebras, in: *Contributions to universal algebra* (Proc. Conf. Szeged, 1975), Coll. Math. Soc. János Bolyai, Vol. 17, North-Holland (Amsterdam, 1977); pp. 503—514.

[16] R. WILLE, *Kongruenzklassengeometrien*, Lecture Notes in Math. 113, Springer (Berlin—Heidelberg—New York, 1970).

DEPARTMENT OF MATHEMATICS
LA TROBE UNIVERSITY
BUNDOORA, VICTORIA, AUSTRALIA 3083

Acta Sci. Math., 51 (1987), 57—71

A decomposition theorem for modular lattices containing an n-diamond

RALPH FREESE[1]

Dedicated to the memory of András Huhn

In the 1930's Von Neumann developed his concept of an n-frame in order to study the coordinatization of complemented modular lattices. In the late 1960's and early 1970's A. P. Huhn revived a variant of this concept, which he called n-diamonds, and used it in his work on modular lattices which were not necessarily complemented. He developed the basic theorems for this concept including the result that n-diamonds (and n-frames) are a projective configuration for the class of modular lattices, [12]. This means that if $f\colon L \twoheadrightarrow M$ is a surjection of modular lattices and M contains an n-diamond then this n-diamond can be pulled back through f to an n-diamond in L.

One of the main themes of modern lattice theory has been the study of lattice varieties. By Birkhoff's theorem, in order to study varieties one needs to understand the operators **H**, **S** and **P** (the closure of classes of algebras under homomorphisms, subalgebras, and direct products, respectively). In the post Jónsson's theorem era of the 1970's, the major unsolved problems on varieties of lattices centered on **H**. It is here that Huhn's result is so useful. Von Neumann showed that associated with each n-frame (and hence each n-diamond) in a modular lattice is a ring. This fact, together with Huhn's projectivity result, has played a crucial role in many of the most important results on modular varieties, certainly in the author's best work.

In this paper we prove the following decomposition theorem, analogous to Fitting's lemma, for finite dimensional modular lattices containing an n-frame. The definitions will be given below.

Theorem 1. *Let L be a finite dimensional modular lattice containing a spanning n-frame, $n \geqq 4$. Then L is a finite direct product of lattices L_i where the ring,*

Received July, 1986, and in revised form October 6, 1986.

[1] This research partially supported by the National Science Foundation.

$R(L_i)$, associated with (the frame in) L_i has prime power characteristic or the field of rational numbers $\mathbf{Q}$ is a subring of $R(L_i)$.

One of the deepest and most important results on modular lattices containing an n-frame is Christian Herrmann's characterization of all subdirectly irreducible modular lattices generated by an n-frame, $n \geqq 4$ [9]. Herrmann's result builds on Huhn's idea of representing automorphisms of frames [13] and the author's result [5] which proves Herrmann's result in the case that the ring associated with the frame has prime characteristic. With the aid of his theorem Herrmann was able to prove the following very powerful result on varieties of modular lattices. Let $\mathcal{M}_0$ denote the variety generated by all subspace lattices of vector spaces over $\mathbf{Q}$.

Theorem 2. (Herrmann [9]). *Every variety of modular lattices which contains $\mathcal{M}_0$ either is not generated by its finite dimensional members or does not have a finite equational basis.*

The following corollary illustrates the power of this theorem. Let $\mathcal{M}$ denote the variety of all modular lattices and let $\mathcal{M}_{\mathrm{f}}$ and $\mathcal{M}_{\mathrm{fd}}$ denote the variety generated by all finite (respectively finite dimensional) modular lattices. Let $\mathcal{A}$ be the variety of all arguesian lattices. (The arguesian law, which is due to Jónsson [14], is stronger than the modular law and related to Desargues law in projective geometry.)

Corollary 3. *$\mathcal{A}$ is not generated by its finite dimensional members. Neither $\mathcal{M}_{\mathrm{f}}$ nor $\mathcal{M}_{\mathrm{fd}}$ is finitely based. Moreover, $\mathcal{M}_{\mathrm{f}} < \mathcal{M}_{\mathrm{fd}} < \mathcal{M}$.*

Since Herrmann's proof of Theorem 2 uses his characterization described above, his proof is quite lengthy. We use the decomposition theorem to give a short proof of his theorem.

The first two sections of this paper give the basic definitions and some results about these concepts. Theorem 1 is proved in the third section. The fourth section proves Herrmann's result, Theorem 2. The fifth section uses a new result of the author to show that the lattices Herrmann used to prove his theorem have a representation by permuting equivalence relations, i.e., a type I representation. The sixth section examines the case of 3-frames. In this case the "ring" associated with the frame may not really be a ring. Nevertheless, an analogue of Theorem 1 can be proved.

1. Preliminaries. We use $+$ and $\cdot$ or juxtaposition to indicate lattice join and meet. An *n-frame* in a lattice L is a subset $\{a_i, c_{ij}\colon i \neq j \text{ and } 1 \leqq i \leqq n\}$ of L such that

$$a_i \cdot \bigvee_{j \neq i} a_j = \bigwedge_{k=1}^{n} a_k, \tag{1}$$

(2) $$a_i \cdot c_{ij} = a_j \cdot c_{ij} = a_i \cdot a_j = \bigwedge_k a_k,$$

(3) $$a_i + c_{ij} = a_j + c_{ij} = a_i + a_j,$$

(4) $$c_{ij} = c_{ji},$$

(5) $$(c_{ij} + c_{jk})(a_i + a_k) = c_{ik}$$

for all distinct i, j, k between 1 and n. A set $\{a_1, \ldots, a_n\}$ is called *independent over* $\bigwedge_k a_k$ if (1) holds. An n-frame in L is called a *spanning n-frame* if $\bigwedge_k a_k = 0_L$ and $\bigvee_k a_k = 1_L$. Let $\{a_i, c_{ij}\}$ be an n-frame, $n \geqq 4$, in a modular lattice L. The ring associated with this frame is, (where we use $\oplus$ and $\otimes$ to denote the addition and multiplication to avoid confusion with the lattice operations)

(6) $$R = \{x \in L\colon\ x + a_2 = a_1 + a_2,\ x \cdot a_2 = a_1 \cdot a_2\},$$

and for $x, y \in R$,

(7) $$x \oplus y = [(x + c_{13})(a_2 + a_3) + (y + a_3)(a_2 + c_{13})](a_1 + a_2),$$

(8) $$x \otimes y = [(x + c_{23})(a_1 + a_3) + (y + c_{13})(a_2 + a_3)](a_1 + a_2).$$

By Theorem 8.4 and Lemma 6.1 of [15] R is a ring with zero a_1 and unit c_{12}. From now on L will denote a lattice containing a fixed spanning n-frame, $n \geqq 4$, and $R(L)$ will denote the ring associate with this n-frame. At the end of the paper the case $n = 3$ will be discussed.

Lemma 1.1. *An element $x \in R(L)$ is invertible if and only if $x + a_1 = a_1 + a_2$ and $x \cdot a_1 = 0$.*

Proof. An elementary proof is given in [6].

An element b in L is called *homogeneous* (*with respect to the frame* $\{a_i, c_{ij}\}$) if $b_i = a_i \cdot b$, $i = 1, \ldots, n$, satisfy $b = \bigvee_i b_i$ and

(9) $$b_j = a_j(b_i + c_{ij}).$$

Whenever b is homogeneous, we shall use the notation $b_i = a_i \cdot b$. The next lemma can be proved with easy calculations.

Lemma 1.2. (i) *Let $k \leqq n$ and suppose that we have an element $b_k \in L$ such that $0 \leqq b_k \leqq a_k$. Let $b_i = a_i(b_k + c_{ik})$, $i \neq k$, and $b = \bigvee_i b_i$. Then b is homogeneous.*

(ii) *If b is homogeneous then $\{a_i + b,\ c_{ij} + b\}$ is an n-frame which spans the interval $1/b$, and $\{a_i \cdot b,\ c_{ij} \cdot b\}$ is an n-frame spanning $b/0$.*

We denote the rings associated with these frames by $R(1/b)$ and $R(b/0)$. More generally, if $b \leqq e$ are both homogeneous, then $\{a_i \cdot e + b,\ c_{ij} \cdot e + b\}$ is a frame which spans e/b. Its ring is denoted $R(e/b)$.

2. Stabilizers. One of the difficulties of these concepts is that if $x \in R(L)$ it does not follow that $x+b \in R(1/b)$. If b is homogeneous we define the *stabilizer of* b, denoted R_b, by

(10) $$R_b = \{x \in R(L)\colon\ b_1 \leqq x+b_2\},$$

where, of course, $b_i = a_i \cdot b$. We say that $x \in R(L)$ is *stable* if it is in R_b for every homogeneous b. (This differs slightly from Herrmann's use of this term in [8].) The next lemmas collect the basic information on stable elements.

Lemma 2.1. *Let L be a modular lattice containing a spanning n-frame and let b be a homogeneous element. Let $x \in R(L)$. Then the following are equivalent:*

(i) $x \in R_b$,
(ii) $x+b \in R(1/b)$,
(iii) $x \cdot b \in R(b/0)$.

Proof. Suppose that x satisfies (i). To show that $x+b \in R(1/b)$ we need to prove that $(x+b)(a_2+b)=b$. Using (i) and the independence of the a_i's we calculate

$$(x+b)(a_2+b) = a_2(x+b)+b = a_2(a_1+a_2)(x+b)+b =$$
$$= a_2\big(x+(a_1+a_2)b\big)+b = a_2(x+b_1+b_2)+b =$$
$$= a_2(x+b_2)+b = a_2 \cdot x+b_2+b = b.$$

Thus (i)$\to$(ii). To see that (i)$\to$(iii) we need to show that $x \cdot b+a_2 \cdot b=b_1+b_2$.

$$x \cdot b+a_2 \cdot b = x \cdot b+b_2 = (x+b_2) \cdot b = (x+b_1+b_2) \cdot b = x \cdot b+b_1+b_2 =$$
$$= x \cdot (a_1+a_2) \cdot b+b_1+b_2 = x \cdot (b_1+b_2)+b_1+b_2 = b_1+b_2.$$

Now if $x \cdot b \in R(b/0)$ then $x \cdot b+b_2=b_1+b_2$. Hence $b_1 \leqq x+b_2$. Hence (iii)$\to$(i). Similarly (ii)$\to$(i).

Lemma 2.2. *R_b is a subring of $R(L)$ closed under taking inverses when they exist. The maps $x \mapsto x+b$ and $x \mapsto x \cdot b$ are ring homomorphisms form R_b to $R(1/b)$ and $R(b/0)$, respectively.*

Proof. By (9) both a_1 and c_{12} (the zero and one of $R(L)$) are in R_b. If $x, y \in R_b$ then using (8) and (9)

$$x \otimes y+b_2 = [(x+c_{23})(a_1+a_3)+(y+c_{13}+b_2)(a_2+a_3)](a_1+a_2) =$$
$$= [(x+c_{23})(a_1+a_3)+(y+c_{13}+b_1+b_2)(a_2+a_3)](a_1+a_2) =$$
$$= [(x+c_{23})(a_1+a_3)+(y+c_{13}+b_1+b_2+b_3)(a_2+a_3)](a_1+a_2) =$$
$$= [(x+c_{23})(a_1+a_3)+b_3+(y+c_{13}+b_1+b_2)(a_2+a_3)](a_1+a_2) \geqq$$
$$\geqq (x+c_{23}+b_3)(a_1+a_3)(a_1+a_2) = (x+c_{23}+b_3+b_2)(a_1+a_3)(a_1+a_2) \geqq b_1.$$

Thus $x\otimes y\in R_b$. Similarly $x\oplus y\in R_b$. If x is invertible in $R(L)$ then a formula for x^{-1} is given in [6]. Using this, one can show that $x^{-1}\in R_b$.

We let $\otimes^b$ denote the multiplication for $R(1/b)$ and $\otimes_b$ for $R(b/0)$. Let $x, y\in R_b$. Then $b_1\leqq b_2+y$ and since $b_i+c_{ij}=b_j+c_{ij}$ by (9), we have,

$$(x+b)\otimes^b(y+b) = [(x+c_{23}+b)(a_1+a_3+b)+(y+c_{13}+b)(a_2+a_3+b)](a_1+a_2+b) =$$
$$= [(x+c_{23}+b_1+b_3)(a_1+a_3)+(y+c_{13}+b)(a_2+a_3)+b](a_1+a_2)+b =$$
$$= [(x+c_{23})(a_1+a_3)+(y+c_{13}+b)(a_2+a_3)+b](a_1+a_2)+b =$$
$$= [(x+c_{23})(a_1+a_3)+(y+c_{13}+b_2)(a_2+a_3)+b_1+b_2](a_1+a_2)+b =$$
$$= [(x+c_{23})(a_1+a_3)+(y+c_{13})(a_2+a_3)](a_1+a_2)+b = x\otimes y+b.$$

These and similar calculations show that $x\mapsto x+b$ and $x\mapsto x\cdot b$ are ring homomorphisms from R_b into $R(1/b)$ and $R(b/0)$.

Notice that this lemma implies that if x is in the prime subring of $R(L)$ (the subring generated by 1) or is the inverse of an element in the prime subring then x is stable.

Notation and motivation. If S is a ring and M is a unitary left S-module then the lattice of submodules, $L(M^n)$, of the module M^n, contains a natural spanning n-frame, namely,

$$a_i = \{(0, \dots, \overset{i\text{th}}{x}, \dots, 0)\colon\ x\in M\},$$

$$c_{ij} = \{(0, \dots, \overset{i\text{th}}{x}, \dots, \overset{j\text{th}}{-x}, \dots, 0)\colon\ x\in M\}.$$

Linear algebraic calculations show that the ring associated with this frame, $R(L(M^n))$, is the endomorphism ring of M. A homogeneous element has the form $\{(x_1, \dots, x_n)\colon\ x_i\in B\}$ for some submodule B of M. The stabilizer of this homogeneous element is the subring of those endomorphisms of M which map B into itself. Simple calculations also show that if $r\in R(L(M^n))$ then $a_1\cdot r=\{(x, 0, \dots, 0)\colon\ xr=0\}$, i.e., $a_1\cdot r$ is the kernel of r (in the first coordinate). Similarly, $a_2(a_1+r)$ is the range of r (in the second coordinate).

For a general modular lattice containing an n-frame, and $x\in R(L)$, there are, by Lemma 1.2, homogeneous elements $b(x)$ and $d(x)$ such that $b(x)_1=a_1\cdot x$ and $d(x)_2=a_2(a_1+x)$. Thus $b(x)$ corresponds to the kernel of x and $d(x)$ to the image.

Lemma 2.3. *Let* $x\in R(L)$ *and let* $b=b(x)$ *and* $d=d(x)$. *Then* $x\in R_b$ *and* $x\in R_d$ *and*

(i) $x+d$ *is the zero element of* $R(1/d)$, *and*

(ii) $x\cdot b$ *is the zero element of* $R(b/0)$.

Proof. Since $x+a_2=a_1+a_2$,

$$x+d_2=x+a_2(x+a_1)=(x+a_2)(x+a_1)=x+a_1(x+a_2)=x+a_1\geqq a_1.$$

Thus $x\in R_d$. Trivially $x\in R_b$. The above calculation shows that $x+d_2=x+a_1$. Hence,

$$a_1+d=a_1+d_2+d=(a_1+a_2)(x+a_1)+d=x+a_1+d=x+d_2+d=x+d.$$

Since a_1+d is the zero element of $R(1/d)$, this proves (i). Again (ii) is trivial; $x\cdot b=b_1$, which is the zero of $R(b/0)$.

3. Proof of Theorem 1.

Lemma 3.1. *If* $u\leqq a_1$ *then*

$$(u+c_{12})a_2=((u+c_{13})a_3+c_{23})a_2.$$

Proof. Let w be the right side of the above equation. Then

$$w+c_{23}+c_{13}=((u+c_{13})a_3+c_{23})(a_2+a_3)+c_{13}=$$
$$=(u+c_{13})a_3+c_{23}+c_{13}=u(a_1+a_3)+c_{23}+c_{13}=u+c_{23}+c_{13}.$$

Meeting both sides with a_1+a_2 gives $w+c_{12}=u+c_{12}$. Thus since $w\leqq a_2$, $w=(w+c_{12})a_2=(u+c_{12})a_2$, as desired.

For $x\in R(L)$ we let x^2 denote $x\otimes x$.

Lemma 3.2. *Let* $x\in R(L)$ *and suppose that* $a_1\cdot x=a_1\cdot x^2$. *Let* $b=b(x)$. *Then*

$$(a_1+b)(x+b)=b.$$

Proof. Since $x\leqq a_1+a_2$, $a_1(x+b)=a_1(a_1+a_2)(x+\bigvee b_i)=a_1(x+b_1+b_2)=b_1+a_1(x+b_2)$. (In the future we shall omit the details of these independence type arguments.) Thus

$$(a_1+b)(x+b)=b+a_1(x+b)=b+a_1(x+b_2)=b+a_1\big(x+a_2(c_{12}+a_1\cdot x)\big).$$

Now we calculate, using the last lemma

$$a_1\cdot x^2=a_1[(x+c_{23})(a_1+a_3)+(x+c_{13})(a_2+a_3)]=$$
$$=a_1(a_1+a_3)[(x+c_{23})(a_1+a_3)+(x+c_{13})(a_2+a_3)]=$$
$$=a_1[(x+c_{23})(a_1+a_3)+a_3(x+c_{13})]=a_1(a_1+a_3)\big(x+c_{23}+a_3(x+c_{13})\big)=$$
$$=a_1\big(x+c_{23}+a_3(a_1+a_3)(x+c_{13})\big)=a_1\big(x+c_{23}+a_3(a_1\cdot x+c_{13})\big)\geqq$$
$$\geqq a_1\big(x+a_2(c_{23}+a_3(a_1\cdot x+c_{13}))\big)=a_1\big(x+a_2(c_{12}+a_1\cdot x)\big).$$

Thus $(a_1+b)(x+b)\leqq b+a_1\cdot x^2=b+a_1\cdot x=b+b_1=b$.

Lemma 3.3. *If L is finite dimensional and $x\in R(L)$ satisfies $a_1\cdot x=0$, then $a_1+x=a_1+a_2$ and thus x is invertible.*

Proof. Suppose that $a_1\cdot x=0$. It is easy to see that we have the following transpositions

$$(a_1+x)/a_1 \searrow x/0 \nearrow (a_1+a_2)/a_2 \searrow c_{12}/0 \nearrow (a_1+a_2)/a_1.$$

Thus the dimension of $(a_1+x)/a_1$ equals that of $(a_1+a_2)/a_1$. Since $a_1+x\leqq a_1+a_2$, this forces equality and the result follows.

Theorem 3.4. *Let L be finite dimensional, $x\in R(L)$ and let $b=b(x)$ and $d=d(x)$ be the elements defined in Section 2. Suppose that x satisfies $a_1\cdot x=a_1\cdot x^2$, then $x+b$ is invertible in $R(1/b)$ and $a_1=b_1+d_1$.*

Proof. That $x+b$ is invertible in $R(1/b)$ follows from Lemmas 1.1, 3.2, and 3.3. Now let $e_1=b_1+d_1$ and let $e=b+d$ be the homogeneous element associated with e_1 (cf. Lemma 1.2). By Lemma 2.3 $x\in R_b$ and R_d. From this it follows that $x\in R_e$. Now since $e\geqq d$, e is a homogeneous element for the frame $\{a_i+d, c_{ij}+d\}$ and by Lemma 2.1 $x+d\in R(1/d)_e$ since $x+d+e=x+e$. By Lemma 2.2 we have three ring homomorphisms, f: $R_d\to R(1/d)$, g: $R(1/d)_e\to R(1/e)$, and h: $R_e\to\to R(1/e)$. Clearly, $g(f(x))=h(x)$. Since $f(x)=x+d$ is the zero element of $R(1/d)$ by Lemma 2.3, $h(x)=x+e$ is the zero element of $R(1/e)$. However, $x+b$ is invertible in $R(1/b)$. By Lemma 2.3 there is a homomorphism of $R(1/b)_e$ into $R(1/e)$ and the image of $x+b$ is $x+b+e=x+e$. Thus $x+e$ is an invertible element of $R(1/e)$. Checking the definition of the ring of a frame one sees that the only way an element of the ring can be both zero and invertible is if the frame is trivial. Thus $e=1$ and thus $a_1\cdot e=a_1=b_1+d_1$, as desired.

Theorem 3.5. *Let L be finite dimensional, $x\in R(L)$ and let $b=b(x)$ and $d=d(x)$ be the elements defined in Section 2. Suppose that x* satisfies *$a_1+x=a_1+x^2$, then $x\cdot d$ is invertible in $R(d/0)$ and $0=b_1\cdot d_1$.*

Proof. Since $x\leqq a_1+a_2$, we have using (9)

$$\begin{aligned}
d_2 &= a_2(a_1+x) = a_2(a_1+x^2) = \\
&= a_2[a_1+(x+c_{13})(a_2+a_3)+(x+c_{23})(a_1+a_3)] = \\
&= a_2[(x+c_{13})(a_2+a_3)+(x+a_1+c_{23})(a_1+a_3)] = \\
&= a_2[(x+c_{13})(a_2+a_3)+a_1+a_3(x+a_1+c_{23})] = \\
&= a_2[(x+c_{13})(a_2+a_3)+a_3(x+a_1+c_{23})] = a_2(x+c_{13}+a_3(x+a_1+c_{23})) = \\
&= a_2[x+c_{13}+a_3(c_{23}+(x+a_1)(a_2+a_3))] = a_2[x+c_{13}+a_3(c_{23}+a_2(x+a_1))] = \\
&= a_2[x+(a_1+a_2)(c_{13}+a_3(c_{23}+a_2(x+a_1)))] = a_2[x+a_1(c_{13}+a_3(c_{23}+a_2(x+a_1)))] = \\
&= a_2[x+a_1(c_{13}+a_3(c_{23}+d_2))] = a_2[x+a_1(c_{13}+d_3)] = a_2(x+d_1).
\end{aligned}$$

Thus $d_2 \leqq x+d_1$ and so $x \cdot d+d_1=d(x+d_1) \geqq d \cdot d_2=d_2$. Hence $x \cdot d+d_1=d_1+d_2$. By an argument similar to the proof to Lemma 3.3, this in turn implies $x \cdot d \cdot d_1= =x \cdot d_1=0$, and thus $x \cdot d$ is invertible in $R(d/0)$ by Lemma 1.1. If we let $e=b \cdot d$ then by the argument of the last theorem, $x \cdot e$ is both the zero and invertible in $R(e/0)$, showing that $e=0$. Hence, $b_1 \cdot d_1=0$, as desired.

It is easy to see that for $x \in R(L)$, $a_1 \cdot x \leqq a_1 \cdot x^2 \leqq a_1 \cdot x^3 \leqq \ldots$, and $a_1+x \geqq \geqq a_1+x^2 \geqq a_1+x^3 \geqq \ldots$. For example, to see the former let $y \in R(L)$. Then

$$a_1 \cdot (x \otimes y) = a_1[(x+c_{23})(a_1+a_3)+(y+c_{13})(a_2+a_3)] =$$

$$= a_1[(x+c_{23})(a_1+a_3)+a_3(y+c_{13})] = a_1(x+c_{23}+a_3(y+c_{13})) \geqq a_1 \cdot x$$

from which $a_1 \cdot x \leqq a_1 \cdot x^2 \leqq \ldots$ follows.

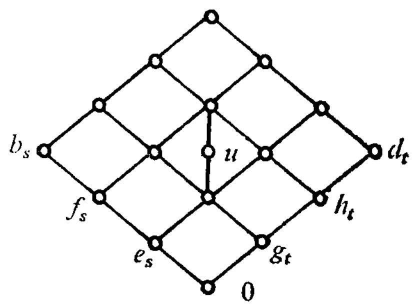
Now if L is finite dimensional there is a k such that $a_1 \cdot x^{2k}=a_1 \cdot x^k$ and $a_1+x^{2k}= =a_1+x^k$. Thus we have the following corollary.

Corollary 3.6. *Let L be a finite dimensional modular lattice containing a spanning n-frame, $n \geqq 4$, and let $x \in R(L)$. Then there are homogeneous elements b and d such that b and d are complements, $x \in R_b \cap R_d$, for some k, x^k is the zero of $R(b/0)$ and x is invertible in $R(d/0)$.*

Proof. As above we choose k such that $a_1 \cdot x^{2k}=a_1 \cdot x^k$ and $a_1+x^{2k}=a_1+x^k$. Let $b=b(x^k)$ and $d=d(x^k)$. Now the result will follow from the previous results once it is shown that $x \in R_b \cap R_d$. To see that $x \in R_b$ we need to show that $b_1 \leqq x+b_2$, i.e., $a_1 \cdot x^k \leqq x+a_2(c_{12}+a_1 \cdot x^k)$. We will actually show $a_1 \cdot x^k \leqq x+a_2(c_{12}+a_1 \cdot x^{k-1})$, which is stronger by the above remarks. We argue by induction on k. It is clear when $k=1$. Now $a_1 \cdot x^k=a_1[(x+c_{23})(a_1+a_3)+(x^{k-1}+c_{13})(a_2+a_3)]$. By the same argument as given in the displayed calculations in the proof of Lemma 3.2, this is equal to $a_1(x+c_{23}+a_3(a_1 \cdot x^{k-1}+c_{13}))$. This equals $a_1(x+a_2(c_{23}+a_3(a_1 \cdot x^{k-1}+c_{13})))$, since $x \leqq a_1+a_2$. By (9) the latter equals $a_1(x+a_2(a_1 \cdot x^{k-1}+c_{12}))$. Thus $a_1 \cdot x^k= =a_1(x+a_2(a_1 \cdot x^{k-1}+c_{12})) \leqq x+a_2(c_{12}+a_1 \cdot x^{k-1})$, as desired. The proof that $x \in R_d$ is similar.

Theorem 3.7. *Let L be a finite dimensional modular lattice containing a spanning n-frame, $n \geqq 4$, and let p be a prime. Then L can be decomposed as $L \cong L_1 \times \times L_2$ in such a way that the characteristic of $R(L_1)$ is a power of p and p is invertible in $R(L_2)$.*

Proof. We view p as an element of $R(L)$. As in the last corollary there is a k such that $a_1+p^{2k}=a_1+p^k$ and $a_1 \cdot p^{2k}=a_1 \cdot p^k$. We again let $b=b(p^k)$ and $d=d(p^k)$. Let $L_1=b/0$ and $L_2=d/0$. By the last result p^k is zero in $R(b/0)=R(L_1)$ and is invertible in $R(d/0)=R(L_2)$. Hence the characteristic of $R(L_1)$ is p^s for some $s \leqq k$, and p is invertible in $R(L_2)$. Also b and d are complements. Since L is modular, this implies that $L_1 \times L_2 \subseteq L$ (this is a "folklore" theorem of lattice theory, see [1] p. 73). In order to show that $L_1 \times L_2 = L$ we need to show that b and d are a *distributive pair*, i.e., for any $u \in L$, b, d, and u generate a distributive sublattice of L (see Theorem 5.2, p. 33 of [15] or 15.9 of [2]). Now we use the following easy result (see Lemma 5.1, p. 36 [15]): if both (b', d) and (b'', d) are distributive pairs and if $b' \cdot d = 0 = b'' \cdot d$, then $(b' + b'', d)$ is a distributive pair and $d(b' + b'') = 0$. Now if (b, d) is not a distributive pair then by repeatedly applying this result there are indices s and t such that (b_s, d_t) is not a distributive pair, i.e., there is a $u \in L$ such that the sublattice generated by b_s, d_t, and u, $\langle b_s, d_t, u \rangle$, is not distributive. Then $\langle b_s, d_t, u(b_s + d_t) \rangle$ is also nondistributive. Hence we may assume that $u \leqq b_s + d_t$. Thus the sublattice generated by u, b_s, and d_t will be a (nondistributive) homomorphic image of the following:



Note that

$$b_s(u + d_t)/b_s \cdot u \nearrow (u + b_s)(u + d_t)/u \searrow d_t(u + b_s)/d_t \cdot u.$$

Since L is finite dimensional we may assume that $b_s \cdot u \prec b_s(u + d_t)$. Let $e_s = u \cdot b_s$, $f_s = b_s(u + d_t)$, $g_t = u \cdot d_t$, and $h_t = d_t(u + b_s)$. We let e be the homogeneous element associated with e_s using Lemma 1.2. We define homogeneous elements f, g, and h in a similar way. Now since $f_i + e \succ e$, f is the join of the atoms above e. This implies that f/e is complemented, see 4.1 of [2]. A complemented modular lattice containing an n-frame, $n \geqq 4$, is isomorphic to the lattices of subspaces, $L(V)$, of an n-dimensional vector space, V, over a skew field F, see 13.4 and 13.5 of [2]. Since $e \leqq b$, and p^k is a stable element of $R(b/0)$, and p^k is zero in $R(b/0)$, the characteristic of F is p. By a similar argument h/g is isomorphic to the lattice of subspaces, $L(U)$, of a vector space, U, over a skew field K in which p^k, and hence p, is invertible.

Since $b \cdot d = 0$ we have that $f/e \nearrow f + g/e + g$ and $h/g \nearrow h + e/e + g$ and $(f + g)(h + e) = e + g$. Thus both $L(V)$ and $L(U)$ can be embedded into $f + h/g + e$. Moreover, since the atoms of $f + h/e + g$ join to $f + h$, $f + h/e + g$ is a complemented modular lattice of length $2n$. Now

$$f_s + e + g/e + g \nearrow f_s + u + e + g/u + e + g = h_t + u + e + g/u + e + g \searrow h_t + e + g/e + g.$$

Since the f_i+e are part of an n-frame, $f_i+e+g/e+g$ is projective to $f_j+e+g/e+g$ for any i and j. Similarly, $h_i+e+g/e+g$ and $h_j+e+g/e+g$ are projective. Hence $f+g$ is the join of pairwise perspective atoms in $f+h/e+g$. Consequently, $f+h/e+g$ is a simple, complemented modular lattices and thus isomorphic the lattice of subspaces of a vector space. But this vector space lattice contains subspace lattices of different characteristics, an impossibility. This contradiction proves the theorem.

To prove Theorem 1 is now easy. Let L be a finite dimensional modular lattice containing an n-frame, $n \geqq 4$. If every prime is invertible in $R(L)$, then $\mathbf{Q}$ is embedded in $R(L)$. If p is not invertible in $R(L)$, then $L \cong L_1 \times L_2$ with p invertible in $R(L_2)$ and the characteristic of $R(L_1)$ a power of p by the last theorem. Now we apply the same procedure to L_2. Since L is finite dimensional, this must stop after finitely many steps and we arrive at the conclusion of the theorem.

4. Herrmann's Theorem. In this section we use Theorem 1 to prove Herrmann's result. Let p be a prime and let $R=\check{\mathbf{Z}}_p$ be the ring of p-adic integers. Recall that the only nonzero ideals of R are p^kR, $k=0, 1, \ldots$. Thus the lattice of submodules of R as a left R-module, $L({}_RR)$, is a descending chain with 0, i.e., the dual of $\omega+1$. Hence $L_1=L({}_RR^n)$ also has the ascending chain condition. If we let a_i be the submodule of R^n generated by $(0, \ldots, 1, \ldots, 0)$, 1 in the i^{th} place, c_{ij} the submodule generated by $(0, \ldots, 1, \cdots, -1, \ldots, 0)$, where the 1 and -1 are in the i^{th} and j^{th} position, then $\{a_i, c_{ij}\}$ is an n-frame in L_1. Now in a modular lattice the relation which identifies a and b if $a+b/a\cdot b$ is finite dimensional is a congruence which we denote here by Θ. Note that $\{a_i/\Theta,\ c_{ij}/\Theta\}$ is a spanning n-frame of L_1/Θ and that a_i/Θ covers 0 in L_1/Θ. As in the last section this implies that $L_1/\Theta \cong L({}_FF^n)$ for some skew field F. Since the operations of $R(L_1)$ are defined from the lattice operations, the homomorphism $L_1 \to L_1/\Theta$ induces a ring homomorphism of $R(L_1)$ into F. It is not hard to see that each member of the frame is the greatest member of its Θ-class. Consequently the only element of $R(L_1)$ which is Θ-equivalent to a_1 is a_1, i.e., the ring homomorphism $R \to F$ is a monomorphism. Hence R is a subring of F. Thus the field of fractions, $\hat{\mathbf{Q}}_p$ of $R=\hat{\mathbf{Z}}_p$, is a subfield of F. (Actually it is not hard to see that $F=\hat{\mathbf{Q}}_p$ and that the homomorphism of L_1 onto L_1/Θ is given by the tensor product: $U \mapsto U \otimes_R \hat{\mathbf{Q}}_p$. This follows from the flatness of $\hat{\mathbf{Q}}_p$ as $\hat{\mathbf{Z}}_p$-module, see 3.32 of [16].) In particular F is uncountable and has characteristic 0.

Since L_1 satisfies the ascending chain condition, each element x of L_1/Θ has a largest inverse image, denoted αx. Thus α is a meet preserving map from L_1/Θ into L_1 mapping the frame in L_1/Θ to the frame of L_1.

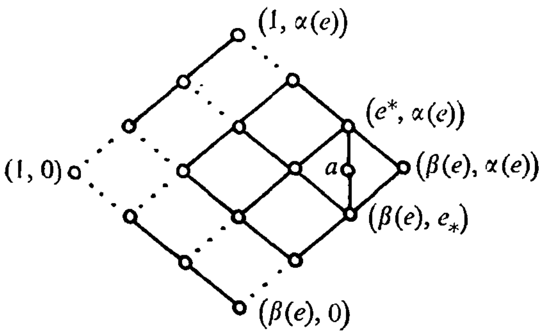
Now let S be the nonmodular lattice obtained from $L({}_\mathbf{Q}\mathbf{Q}^n)$ by adjoining an extra element e which is between 0 and 1 and a complement of all other elements.

Since $\mathbf{Q}$ is embedded in F, $L({}_{\mathbf{Q}}\mathbf{Q}^n)$ is a sublattice of $L({}_F F^n)$ in a natural way. (This map sends a subspace U to $U\otimes_{\mathbf{Q}} F$, which is just the F-subspace generated by U.) Extend this map to S by mapping e to a point which is on no rational hyperplane. For example, e can be sent to a one dimensional subspace spanned by a vector in F^n whose coordinates are linearly independent over $\mathbf{Q}$. Combining this map with α we obtain a meet embedding of S into L_1, which is canonical on the frames. We also use α to denote this map. Thus α maps e to a rank 1 free submodule, and hence $\alpha(e)/0$ is dually isomorphic to the ordinal $\omega+1$.

Let L_0 be the lattice which is dual to L_1, except we use the prime q in place of p. (L_0 may be taken to be the lattice of subgroups of the direct product of n copies of the Prüfer group $\mathbf{Z}_{q^\infty}$.) Then there is a join embedding β of S into L_0 such that $1/\beta(e)$ is isomorphic to $\omega+1$. Let

$$A_{pq} = \{(u, v)\in L_0\times L_1\colon \exists x\in S,\ \beta(x) \leqq u,\ v \leqq \alpha(x)\}.$$

It is easy to check that this is a sublattice which contains the spanning frame $\{(a_i, a_i), (c_{ij}, c_{ij})\}$. By the above remarks the interval $(1, \alpha(e))/(\beta(e), 0)$ is isomorphic to $(\omega+1)\times(\omega+1)^d$. We let $e^*\in L_0$ denote the upper cover of $\beta(e)$ and $e_*\in L_1$ the lower cover of $\alpha(e)$. We let L_{pq} be the lattice obtained from A_{pq} by adjoining a new element a so that $(e^*, \alpha(e))/(\beta(e), e_*)$ is isomorphic to M_3. Since $(\beta(e), \alpha(e))$ is both join and meet irreducible in A_{pq}, it is easy to see that L_{pq} is a modular lattice. The interval $(1, \alpha(e))/(\beta(e), 0)$ of L_{pq} is drawn below where the solid lines indicate coverings.



Now $R(L_1)\cong\hat{\mathbf{Z}}_p$ and simple linear algebraic calculations show that this isomorphism is given by $r\mapsto\{(-x, rx, 0, \dots, 0)\colon x\in\hat{\mathbf{Z}}_p\}$. Below we identify r and this submodule. Again by linear algebraic calculations we have that $a_1\cdot r=\{(y, 0, \dots, 0)\colon ry=0\}$. In our case $R(L_1)\cong\hat{\mathbf{Z}}_p$ is an integral domain. Hence we have $a_1\cdot r=0$ for each $r\in\hat{\mathbf{Z}}_p$ except $r=0$. Similarly, we have that $(a_1+p)a_2=\{(0, px, 0, \dots, 0)\colon x\in\hat{\mathbf{Z}}_p\}$. Since $p\hat{\mathbf{Z}}_p$ is the unique maximal ideal of $\hat{\mathbf{Z}}_p$, $(a_1+p)a_2$ is the unique lower cover of a_2 in L_1.

Recall that L_0 is isomorphic to the lattice of subgroups of the direct product of n copies of $\mathbf{Z}_{q^\infty}$. In general the ring associated with the direct product of n copies of a module is the endomorphism ring of the module. Thus in our case $R(L_0)\cong \cong \mathrm{End}\,(\mathbf{Z}_{q^\infty})\cong \hat{\mathbf{Z}}_q$. With the aid of these facts, it is not hard to verify that in L_0 $a_1+r=a_1+a_2$ for all $r\in R(L_0)$ except $r=0_R=a_1$, and that $a_1\cdot q$ is the unique atom below a_1. It follows that in L_{pq} $(a_1, a_1)\cdot(q, q)=(a_1\cdot q, 0)\succ 0$, and $[(a_1, a_1)+(p, p)]\cdot(a_2, a_2)=(a_2, (a_1+p)a_2)\prec(a_2, a_2)$. Now in L_0 $a_1\cdot\beta(e)=0$ so that $a_1\cdot q+\beta(e)\succ\beta(e)$. Hence in L_{pq}

$$(a_1\cdot q, 0)/(0, 0)\nearrow(e^*, \alpha(e))/(\beta(e), \alpha(e)).$$

Similarly,

$$(\beta(e), \alpha(e))/(\beta(e), e_*)\nearrow(1, 1)/(1, a_1+p+a_3+\ldots+a_n)\searrow(a_2, a_2)/(a_2, (a_2(a_1+p))).$$

Thus in L_{pq} $(a_2, a_2)/(a_2, a_2(a_1+p))$ and $(a_1\cdot q, 0)/(0, 0)$ are projective prime quotients.

We will show that L_{pq} is not in $\mathcal{M}_{\mathrm{fd}}$. Suppose that $L_{pq}\in\mathcal{M}_{\mathrm{fd}}$. Then L_{pq} is a homomorphic image of a lattice M which is residually finite dimensional. By Huhn's theorem, [12], M has a frame $\{a_i, c_{ij}\}$, which we may assume spans M, mapping onto the frame $\{(a_i, a_i), (c_{ij}, c_{ij})\}$ in L_{pq}. By an easy application of Dedekind's transposition principle, we have, in M, that $a_1q/0$ and $a_2/a_2(a_1+p)$ have nontrivial subquotients which are projective. Thus there are elements $b_1, c_1, f_2, g_2\in M$ such that $0\leqq c_1<b_1\leqq a_1\cdot q$ and $a_2(a_1+p)\leqq g_2<f_2\leqq a_2$ and b_1/c_1 and f_2/g_2 are projective. Let $b, c, f, g\in M$ be the homogeneous elements associated with b_1, c_1, f_2, g_2, see Lemma 1.2. Since M is residually finite dimensional, there is a homomorphism $\psi\colon M\to K$ with K finite dimensional such that $\psi(b_1)\neq\psi(c_1)$. By Theorem 3.7 $K\cong K_1\times K_2$ where $R(K_1)$ has characteristic p^i for some i, and p is invertible in $R(K_2)$. Let $\pi_i\colon K\to K_i$, $i=1, 2$, be the projection homomorphisms.

Since q is in the subring of $R(M)$ generated by 1, it is stable. Thus by Lemma 2.2, q in $R(b/0)$ is the element $q\cdot b$. But since $b_1\leqq a_1\cdot q\leqq q$, $q\cdot b=q(b_1+b_2)=b_1+q\cdot b_2= =b_1$. Thus $R(b/0)$, and hence $R(b/c)$, has characteristic q. Since $a_2(a_1+p)\leqq g_2$, we have, by joining a_1 to both sides, $p\leqq a_1+p\leqq(a_1+a_2)(a_1+p)\leqq a_1+g$. Hence $p+g\leqq a_1+g$, which implies that $p=0$ in $R(1/g)$. Thus $p=0$ in $R(f/g)$, again by Lemma 2.2.

It follows that in K_1, $R(\pi_1\psi(b)/\pi_1\psi(c))$ satisfies $q=0$. But $R(K_1)$ satisfies $p^i=0$ and thus $R(\pi_1\psi(b)/\pi_1\psi(c))$ also satisfies $p^i=0$ by Lemma 2.2. Since p and q are relatively prime, this ring must satisfy $1=0$, i.e., $\pi_1\psi(b)=\pi_1\psi(c)$, so that $(\psi(b), \psi(c))\in\ker\pi_1$. Hence $(\psi(b_1), \psi(c_1))\in\ker\pi_1$. Similarly $(\psi(f_2), \psi(g_2))\in\ker\pi_2$. But $\psi(b_1)/\psi(c_1)$ projects to $\psi(f_2)/\psi(g_2)$. Thus $(\psi(b_1), \psi(c_1))\in\ker\pi_1\cap\ker\pi_2=0$. Hence $\psi(b_1)=\psi(c_1)$, a contradiction. Hence $L_{pq}\notin\mathcal{M}_{\mathrm{fd}}$, as claimed.

Let p^+ be the first prime after p. The next step in the proof is to show that any nonprincipal ultraproduct of $\{L_{pp^+}\colon p \text{ a prime}\}$ lies in $\mathcal{M}_0$. This is a fairly standard argument and we shall only sketch it. HERRMANN's original proof [9] contains more details.

Let $L=(\Pi_p L_{pp^+})/\mathcal{U}$ be a nonprincipal ultraproduct of $\{L_{pp^+}\}$. The corresponding ultraproduct of rings $R=(\Pi_p \hat{\mathbf{Z}}_p)/\mathcal{U}$ has characteristic 0 and every prime is invertible, since these facts hold in $\hat{\mathbf{Z}}_p$ for almost all p. Hence $\mathbf{Q}$ is a subring of R. Now the ultraproduct $(\Pi_p L(\hat{\mathbf{Z}}_p^n))/\mathcal{U}$ is a lattice of submodules of a module over R. Since $\mathbf{Q}\subseteqq R$, this may be viewed as a module over $\mathbf{Q}$. Hence this lattice can be embedded into the lattice of subspaces of a vector space over $\mathbf{Q}$. Let $A= =(\Pi_p A_{pp^+})/\mathcal{U}$. Then A can be embedded into the direct product $L(V_0)\times L(V_1)$ of vector space lattices over $\mathbf{Q}$. Now A is just L with the element $(\Pi a)/\mathcal{U}$ removed. Also in L and in A we have $(\Pi(\beta(e), e_*))/\mathcal{U}\prec(\Pi(\beta(e),\alpha(e)))/\mathcal{U}\prec(\Pi(e^*, a(e)))/\mathcal{U}$. By changing V_0 and V_1 we may assume that $(e^*,\alpha(e))/(\beta(e),\alpha(e))$ and $(\beta(e),\alpha(e))/(\beta(e), e_*)$ have the same dimension. Now $L(V_0)\times L(V_1)$ is a sublattice of $L(V_0\times V_1)$ and L can be embedded into this lattice by choosing a to be a common complement of (e^*, e_*) and $(\beta(e),\alpha(e))$ in $(e^*,\alpha(e))/(\beta(e), e_*)$. Thus $L\in\mathcal{M}_0$.

Now the proof of Theorem 2 can easily be completed. If $\mathcal{K}$ is a finitely based variety with $\mathcal{M}_0\subseteqq\mathcal{K}$ then the ultraproduct $(\Pi_p L_{pp^+})/\mathcal{U}$ lies in $\mathcal{K}$. Since $\mathcal{K}$ is finitely based there must be a prime p such that $L_{pp^+}\in\mathcal{K}$. Since $L_{pp^+}\notin\mathcal{M}_{\mathrm{fd}}$, $\mathcal{K}$ is not generated by its finite dimensional members.

For Corollary 3, the fact that $\mathcal{M}_{\mathrm{f}}\neq\mathcal{M}_{\mathrm{fd}}$ is a result of FREESE [4]. The rest of the corollary follows from the fact that $\mathcal{M}_0\subseteqq\mathcal{M}_{\mathrm{f}}$, which is proved by HERRMANN and HUHN in [11].

5. Type I representations. It follows from the results of FREESE, HERRMANN and HUHN [7] that if $\mathcal{V}$ is a variety of algebras all of whose congruences are modular then L_{pq} is not in the variety generated by the congruence lattices of the algebras in $\mathcal{V}$. Indeed, in the last proof we showed that $(a_2, a_2)/(a_2, a_2(a_1+p))$ and $(a_1\cdot q, 0)/(0, 0)$ are projective prime quotients in L_{pq}. Let b be the homogeneous element of L_0 with $b_1=a_1\cdot q$ and let d be the homogeneous element of L_1 with $d_2=a_2(a_1+p)$, see the notation before Lemma 2.3. It follows from Lemma 2.3 that $R(b/0)$ has characteristic q in L_0 and $R(1/d)$ has characteristic p in L_1. Thus in L_{pq} the ring of the frame in $(b, 0)/(0, 0)$ has characteristic q and the ring of the frame $(1, 1)/(1, d)$ has characteristic p. The projectivity above shows that the quotient $(b_1, 0)/(0, 0)$ in the first frame is projective to $(1, d+a_1)/(1, d)$ in the second. Now Proposition 2 of [7] shows that this situation cannot occur in a modular congruence variety. Thus L_{pq} cannot be in any modular congruence variety.

In light of the above result it is of interest to decide if L_{pq} has a representation as a lattice of permuting equivalence relations (known as a *type I representation*). The following theorem of the author shows that it does have such a representation. The proof of this theorem will appear elsewhere.

Theorem 5.1. *Let* L *be a modular lattice containing an element* a *which is both join and meet irreducible. If the sublattice* $L-\{a\}$ *has a type I representation then* L *has such a representation.*

Summarizing these results we have:

Corollary 5.2. *The lattices* L_{pq} *all have a type I representation. If* $p\neq q$ *and* $\mathcal{K}$ *is a variety of algebras all of whose congruences are modular, then* L_{pq} *is not in the lattice variety generated by the congruences lattices of the members of* $\mathcal{K}$.

6. 3-frames. Throughout the previous sections we dealt with n-frames where n was at least 4. In this section we show that an analogue of Theorem 3.7 holds for $n=3$. The definition of the ring R determined by a frame with $n=3$ given in (6) makes perfect sence. Moreover addition and multiplication, as given in (7) and (8), are well-defined. However, it is not true that $(R, \oplus, \otimes)$ satisfies the ring axioms, as the lattices associated with non-Desarguesian projective planes show. In particular neither operation is necessarily associative. We will call a term in $\oplus$ and $\otimes$ and the constant 1 and no variables a *prime term* if its evaluation in $\mathbf{Z}$ is a prime. Thus $[(1\oplus 1)\otimes(1\oplus 1)]\oplus 1$ is a prime term. By a *prime* in R we will mean the evaluation of a prime term in R. By a *power* of $x\in R$ we mean the evaluation of some term in only $\otimes$ and the variable x.

Theorem 6.1. *Let* L *be a finite dimensional modular lattice containing a spanning 3-frame. Let* $\langle R;, \oplus, \otimes\rangle$ *be the algebraic structure defined by* (6), (7) *and* (8), *and let* p *be a prime in* R. *Then* L *can be decomposed as* $L\cong L_1\times L_2$ *in such a way that in* $R(L_1)$ *some power of* p *is zero and* p *is invertible in* $R(L_2)$.

Proof. The proof is essentially the same as Theorem 3.7. For the most part one simply notes that the proofs work for $n=3$. There are two places where some care is necessary. Define the symmetric power, $x^{[n]}$, of $x\in R$ by $x^{[1]}=x$, and $x^{[n+1]}=x^{[n]}\otimes x^{[n]}=(x^{[n]})^2$. Now before Corollary 3.6 we showed that $a_1\cdot x\leqq a_1(x\otimes y)$, for $x, y\in R$. A similar argument shows that $a_1+(x\otimes y)\leqq a_1+y$. From this we see that $a_1\cdot x\leqq a_1\cdot x^{[2]}\leqq a_1\cdot x^{[3]}\leqq\dots$ and $a_1+x\geqq a_1+x^{[2]}\geqq a_1+x^{[3]}\geqq\dots$. Hence, by the finite dimensionality of L, it follows that some symmetric power y of x satisfies $a_1\cdot y=a_1\cdot y^2$ and $a_1+y=a_1+y^2$.

The other place that requires care is the proof, in Theorem 3.7, that (b, d) is a distributive pair. This required a vector space argument. However, the proof showed that if (b, d) failed to be a distributive pair then L contained a simple complemented sublattice of dimension $2n=6$. Since $6>4$ the classic coordinatization theorem (see 13.4 and 13.5 of [2]) shows that this sublattice is isomorphic to the lattice of subspaces of a vector space. Moreover the proof of Theorem 3.7 shows that this sublattice will have two three-dimensional sublattices. The "rings" determined by the frames of these three dimensional intervals must be real rings because they lie inside a vector space lattice. In one of the rings, a power of p is zero and in the other, it is invertible. This is of course impossible in a vector space lattice.

References

[1] G. BIRKHOFF, *Lattice theory*, 3rd edition, Colloq. Publ., vol. 25, Amer. Math. Soc. (Providence, 1967).

[2] P. CRAWLEY and F. P. DILWORTH, *Algebraic Theory of Lattices*, Prentice-Hall (Englewood Cliffs, N. J. 1973).

[3] R. FREESE, Planar sublattices of $FM(4)$, *Algebra Universalis*, **6** (1976), 69—72.

[4] R. FREESE, The variety of modular lattices is not generated by its finite members, *Trans. Amer. Math. Soc.*, **255** (1979), 277—300.

[5] R. FREESE, Projective geometries as projective modular lattices, *Trans. Amer. Math. Soc.*, **251** (1979), 329—342.

[6] R. FREESE, Free modular lattices, *Trans. Amer. Math. Soc.*, **261** (1980), 81—91.

[7] R. FREESE, C. HERRMANN and A. HUHN, On some identities valid in modular congruence varieties, *Algebra Universalis*, **12** (1981), 322—334.

[8] C. HERRMANN, On the word problem for the modular lattice with four free generators, *Math. Ann.*, **265** (1983), 513—527.

[9] C. HERRMANN, On the arithmetic of projective coordinate systems, *Trans. Amer. Math. Soc.*, **284** (1984), 759—785.

[10] C. HERRMANN and A. P. HUHN, Lattices of normal subgroups which are generated by frames, in: *Lattice Theory* (Proc. Conf. Szeged, 1974), Colloq. Math. Society János Bolyai, vol, 14, North-Holland (Amsterdam, 1976); pp. 97—136.

[11] C. HERRMANN and A. P. HUHN, Zum Wortproblem für freie Untermodulverbände, *Arch. Math. (Basel)*, **26** (1975), 449—453.

[12] A. P. HUHN, Schwach distributive Verbände. I, *Acta Sci. Math.*, **33** (1972), 297—305.

[13] A. P. HUHN, On G. Grätzer's problem concerning automorphisms of finitely presented lattices, *Algebra Universalis*, **5** (1975), 65—71.

[14] B. JÓNSSON, Modular lattices and Desargues' theorem, *Math. Scand.*, **2** (1954), 295—314.

[15] J. VON NEUMANN, *Continuous Geometry*, Princeton University Press (Princeton, N. J., 1960).

[16] J. ROTMAN, *Notes on Homological Algebra*, Van Nostrand (New York, 1970).

DEPARTMENT OF MATEMATICS
THE UNIVERSITY OF HAWAII
2565 THE MALL
HONOLULU, HAWAII 96822, U.S.A.

Acta Sci. Math., 51 (1987), 73—80

The lattice variety $\mathbf{D} \circ \mathbf{D}$

GEORGE GRÄTZER and DAVID KELLY*

Dedicated to the memory of András P. Huhn

Section 1. Introduction. Let $\mathbf{V}$ and $\mathbf{W}$ be varieties of lattices. The product of $\mathbf{V}$ and $\mathbf{W}$, denoted by $\mathbf{V} \circ \mathbf{W}$, consists of all lattices L for which there is a congruence relation Θ such that every congruence class of Θ (as a lattice) is in $\mathbf{V}$ and L/Θ is in $\mathbf{W}$.

In this paper, we investigate in detail the class $\mathbf{D}^2 = \mathbf{D} \circ \mathbf{D}$. This class first appeared in a paper of S. V. POLIN [9]. $\mathbf{D}^2$ is a curious class. Usually, one defines a class of algebras and aims at obtaining a structure theorem, while $\mathbf{D}^2$ is defined *via* a structure theorem: members of $\mathbf{D}^2$ are formed from distributive lattices over another distributive lattice.

In Section 2 we exhibit some lattices in $\mathbf{D}^2$. We describe a method to construct lattices freely generated by a poset over $\mathbf{D}^2$; we apply this (Theorem 1, Figure 1) to obtain the free product over $\mathbf{D}^2$ of the one-element and the four-element chain, and (Theorem 2) the free lattice over $\mathbf{D}^2$ generated by the six-element partially ordered set H (see Figures 2 and 3). An example shows (Theorem 3, Figure 4) that $\mathbf{D}^2$ is not locally finite.

In Section 3 we verify the most important property of $\mathbf{D}^2$: it is a variety. This result is a special case of the following result (Theorem 4): Let $\mathbf{V}$ be a lattice variety closed under gluing; then $\mathbf{V} \circ \mathbf{D}$ is a variety. In particular, $\mathbf{D}^2$ is a variety. As a corollary of this theorem, we get that there are continuumly many pairs of varieties whose product is a variety again.

While most known lattice varieties are either modular (contained in $\mathbf{M}$, the variety of modular lattices) or of small height (their height in the lattice of lattice varieties is 4 or less), $\mathbf{D}^2$ is neither. We show that $\mathbf{D}^2$ has large height (Theorem 5): There are continuumly many varieties contained in $\mathbf{D}^2$. Also, $\mathbf{D}^2$ is very far from $\mathbf{L}$ (the variety of all lattices): there are continuumly many varieties containing $\mathbf{D}^2$. Finally, $\mathbf{D}^2$ is almost disjoint from $\mathbf{M}$: $\mathbf{D}^2 \cap \mathbf{M} = \mathbf{D}$.

Received May 4, 1986.

* The research of both authors is supported by the NSERC of Canada.

The results reported in this paper were discovered in the late seventies. There are some newer results (1982—1985) of the type covered in Section 3. Firstly, there is the result of R. McKenzie (see R. McKenzie and D. Hobby [9]) that $\mathbf{D}\circ\mathbf{V}$ and $\mathbf{M}\circ\mathbf{V}$ are always varieties. The corollary of the main result of Section 3 also follows from McKenzie's result. The result of T. Harrison [6] shows that McKenzie's result is best possible: if $\mathbf{W}$ is a non-modular lattice variety with the property that $\mathbf{W}\circ\mathbf{V}$ is a variety for any given non-modular lattice variety $\mathbf{V}$, then $\mathbf{W}=\mathbf{L}$, the variety of all lattices.

For the basic facts concerning products of lattice varieties, we refer to our paper [4]. For the basic concepts and notation, the reader is referred to [2].

Section 2. Examples. Although it may seem rather restrictive to require a lattice to be in $\mathbf{D}^2$, there are surprisingly many lattices in $\mathbf{D}^2$.

Let us start with small varieties. Obviously, N_5 (the five-element nonmodular lattice) is in $\mathbf{D}^2$, hence $\mathbf{N}_5$ (the variety generated by N_5) is contained in $\mathbf{D}^2$. $\mathbf{N}_5$ has 16 covers (B. Jónsson and I. Rival [7]; 15 of them are generated by the lattices of Figures 3—11 in Section V.2 of [2] and their duals; the 16th is $\mathbf{N}_5\vee\mathbf{M}_3$). All but two are contained in $\mathbf{D}^2$. The exceptions are the (self-dual) variety generated by Figure 11 and $\mathbf{N}_5\vee\mathbf{M}_3$. M_3 does not belong to $\mathbf{D}^2$ because it is simple and it does not belong to $\mathbf{D}$. In fact, the only simple lattice in $\mathbf{D}^2$ is the two-element chain. Which modular lattices belong to $\mathbf{D}^2$? A modular lattice is non-distributive iff it contains M_3; hence, a modular lattice belongs to $\mathbf{D}^2$ iff it is distributive.

It is easy to check whether a lattice belongs to $\mathbf{D}^2$. For a lattice variety $\mathbf{V}$, and a lattice L, let $\Theta(L, \mathbf{V})$ be the smallest congruence relation on L such that $L/\Theta(L, \mathbf{V})$ is in $\mathbf{V}$. Now, L belongs to $\mathbf{D}^2$ iff all $\Theta(L, \mathbf{D})$ congruence classes are distributive. The "if" parts is obvious. Conversely, if L belongs to $\mathbf{D}^2$ by virtue of the congruence relation Θ, then $\Theta\geqq\Theta(L, \mathbf{D})$; since all Θ classes are distributive, so are the $\Theta(L, \mathbf{D})$ classes. $\Theta(L, \mathbf{D})$ can be described as follows: it is the join of all principal congruence relations $\Theta(a, b)$, where $a<b$ is a "violation of the distributive identity"; that is, there are $x, y, z\in L$ such that $(x\wedge y)\vee z=a$ and $(x\vee z)\wedge(y\vee z)=b$.

Using this, we can find the largest homomorphic image of a lattice that belongs to $\mathbf{D}^2$. Indeed, for a lattice L, first form $\Theta(L, \mathbf{D})$. Then form the join Φ of all $\Theta(a, b)$, where $a<b$ is a violation of the distributive law in some $\Theta(L, \mathbf{D})$ congruence class. Obviously, $\Phi\leqq\Theta(L, \mathbf{D})$; in L/Φ, $\Theta(L, \mathbf{D})/\Phi$ has distributive congruence classes, and $\Theta(L, \mathbf{D})/\Phi=\Theta(L/\Phi, \mathbf{D})$. Hence Φ is the same as $\Theta(L, \mathbf{D}^2)$. (We can describe similarly the congruence $\Theta(L, \mathbf{V}\circ\mathbf{W})$ for arbitrary lattice varieties $\mathbf{V}$ and $\mathbf{W}$.)

We apply this observation to determine some lattices freely generated by partially ordered sets over $\mathbf{D}^2$. Let C_n denote the n-element chain, and $A*B$ the free product of A and B. $L=C_2*C_2$ is in $\mathbf{D}^2$; so the $\mathbf{D}^2$-free product of C_2 and C_2 is L (see Figure 6 of Section VI.1 of [2]). However, the free product of C_1 and C_4

(see Figure 7 of Section VI.1 of [2]) does not lie in $\mathbf{D}^2$. Applying the construction of $\Phi=\Theta(L, \mathbf{D}^2)$ to this lattice $L=C_1 * C_4$ we obtain the lattice of Figure 1.

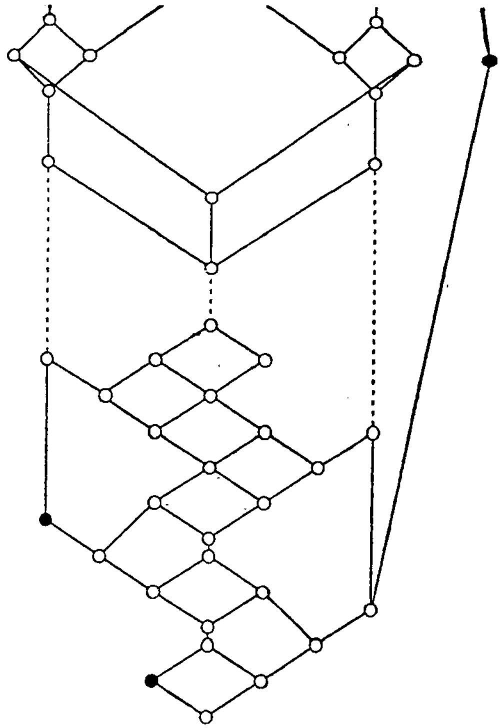


Figure 1

Theorem 1. *The lattice of Figure 1 is the* $\mathbf{D}^2$*-free product of* C_1 *and* C_4.

The free lattice L over the partially ordered set H (see Figure 2) plays an important role in [11] (see also [5]). Applying the method described above to this lattice L, we obtain the lattice freely generated by H over $\mathbf{D}^2$; see Figure 3.

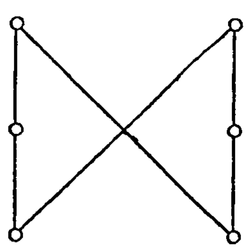


Figure 2

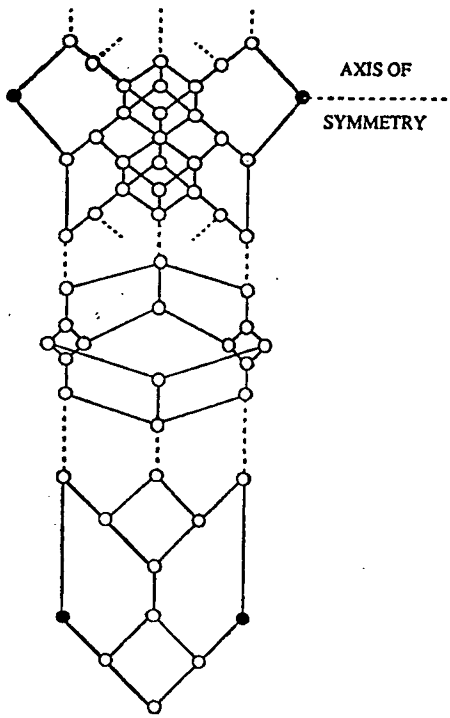


Figure 3

Theorem 2. *The lattice of Figure* 3 *is the* $\mathbf{D}^2$*-free lattice over* H.

Our final example of a lattice in $\mathbf{D}^2$ is Figure 4. Since this is a 3-generated infinite lattice, we conclude that:

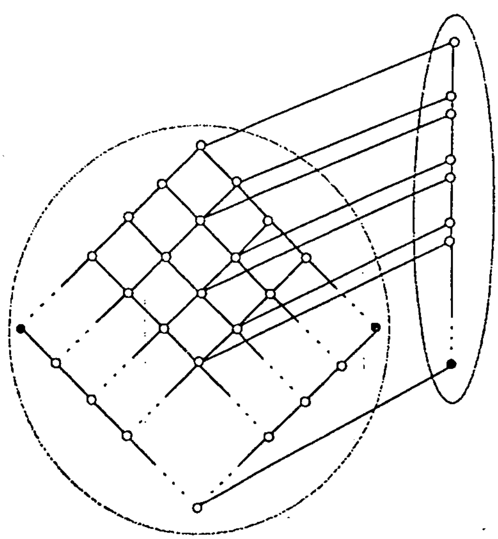


Figure 4

Theorem 3. $\mathbf{D}^2$ *is not locally finite.*

Section 3. $\mathbf{D}^2$ is a variety. We proved in [4] that if $\mathbf{V} \circ \mathbf{D}$ is a variety, then $\mathbf{V} \circ \mathbf{D}$ is closed under gluing. In this section we prove the following theorem:

Theorem 4. *Let* $\mathbf{V}$ *be a lattice variety. If* $\mathbf{V}$ *is closed under gluing, then* $\mathbf{V} \circ \mathbf{D}$ *is a variety.*

Proof. Let $\mathbf{V}$ be a lattice variety closed under gluing, let $L \in \mathbf{V} \circ \mathbf{D}$, and let Φ be a congruence relation of L. Given such an L, there is a smallest congruence relation Θ establishing that $L \in \mathbf{V} \circ \mathbf{D}$. To show that $\mathbf{V} \circ \mathbf{D}$ is a variety it is sufficient to show that for all choices of L and Φ, $L/\Phi \in \mathbf{V} \circ \mathbf{D}$. Since $\mathbf{V} \circ \mathbf{D}$ is a quasi-variety, L/Φ belongs to it iff all finitely generated sublattices of L/Φ belong. Hence we can assume that L is finitely generated. Therefore, L/Φ is a finite distributive lattice. In [4] we have observed that it is sufficient to prove that $L/\Phi \in \mathbf{V} \circ \mathbf{D}$ for Φ satisfying $\Theta \wedge \Phi = \omega$.

Since L/Θ is finite, Φ can be written as the join of n congruences of the form $\Theta(a, b)$ (called *minimal*), where $[a]\Theta$ covers $[b]\Theta$ in L/Θ. We prove that $L/\Phi \in \mathbf{V} \circ \mathbf{D}$ by induction on n. Let $\Phi = \Phi' \vee \Theta(a, b)$. Since L/Φ is isomorphic to $(L/\Phi')/(\Phi/\Phi')$, and Φ/Φ' is minimal in L/Φ', we can assume without loss of generality that $\Phi = \Theta(a, b)$ for such a pair a, b.

We claim that $L/\Phi \in \mathbf{V} \circ \mathbf{D}$ is established by the congruence relation $(\Theta \vee \Phi)/\Phi$. By the Second Isomorphism Theorem (see [2]) $(L/\Phi)/(\Theta \vee \Phi/\Phi)$ is a homomorphic image of L/Θ, hence this lattice is distributive. The behavior of a $\Theta(u, v)$, u covers v, in a distributive lattice is well known (see [2], Chapter II); in particular, every congruence class is a singleton or a covering pair. Each congruence class of L/Φ modulo $\Theta \vee \Phi/\Phi$ lies in $\mathbf{V}$ because it is either isomorphic to a congruence class of L modulo Θ or it is isomorphic to a lattice described in the following lemma.

Lemma 1. *Let* K *be a lattice, and let* $\mathbf{V}$ *be a lattice variety closed under gluing. Let* Θ *be a congruence relation on* K *with two congruence classes which as lattices are in* $\mathbf{V}$. *Let* Φ *be a congruence relation on* K *satisfying* $\Theta \wedge \Phi = \omega$. *Then* $K/\Phi \in \mathbf{V}$.

Proof. Let A and B be the congruence classes of K, with A the zero of L/Θ. Let D be the set of those elements of A that are congruent to some element of B modulo Φ. Let I be the set of those elements of B that are congruent to some element of A. We claim that D is a dual ideal of A, and I is an ideal of B.

Let $a_1, a_2 \in D$. There are $b_1, b_2 \in B$ satisfying $a_1 \equiv b_1(\Phi)$ and $a_2 \equiv b_2(\Phi)$. Then

$$a_1 \wedge a_2 \equiv b_1 \wedge b_2(\Phi).$$

Since $b_1 \wedge b_2 \in B$, we conclude that $a_1 \wedge a_2 \in D$. Also, if $a \in D$ and $x \in A$, then there is a $b \in B$ satisfying $a \equiv b(\Phi)$. Hence,

$$a \vee x \equiv b \vee x(\Phi)$$

and $b \vee x \in B$, verifying that D is a dual ideal.

Similarly, I is an ideal of B.

Now if $a \in D$, then there is a unique $b \in I$ satisfying $a \equiv b(\Phi)$ (otherwise, $\Theta \wedge \Phi = \omega$ would be contradicted). Thus, we have a mapping φ from D to I. It is easy to verify that φ is an isomorphism. Moreover, it is clear that the Φ classes with more that one element are exactly: $\{a, a\varphi\}$, $a \in D$. Thus, A and B glued over I and D is isomorphic to K/Φ, and hence is in $\mathbf{V}$, as claimed.

Section 4. Subvarieties. In this section, we construct continuumly many distinct subvarieties of $\mathbf{D}^2$.

Let A be an atomic Boolean lattice, $|A| \geqq 8$. We construct the lattices $K(A)$ and $L(A)$ as follows (see Figure 5). We take a disjoint copy A' of A. The zero and

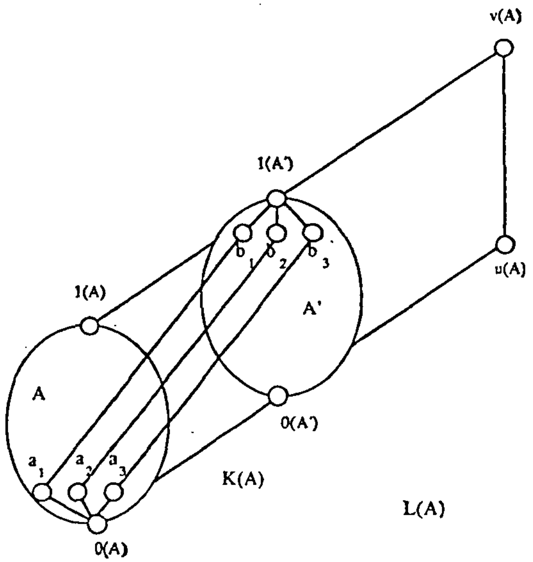


Figure 5

unit of A and A' are denoted by $0(A), 1(A), 0(A'), 1(A')$, respectively. Let $a_1(A), a_2(A), \ldots$ be the atoms of A, and $d_1(A'), d_2(A'), \ldots$ be the dual atoms of A'; if it is clear from the context, we may write a_i for $a_i(A)$ and d_i for $d_i(A')$. $K(A)$ is defined on $A \cup A'$; A and A' are subposets of $K(A)$; for $x \in A$ and $y \in A'$, $x < y$ iff $x = 0(A)$, or $y = 1(A')$, or $x = a_i(A)$, $y = b_i(A')$ for some i; for $x \in A$ and

$y \in A'$, $x > y$ never holds. $L(A)$ is defined on $K(A)$ and two new elements: $u(A)$ and $v(A)$. $v(A)$ is the unit element of $L(A)$; $0(A), 0(A') < u(A)$; $u(A)$ is only comparable to $0(A)$, $0(A')$ and $v(A)$.

Lemma 2. *Let A be an atomic Boolean lattice. Then $K(A)$ and $L(A)$ are lattices. $K(A)$ is a sublattice of $L(A)$. Both lattices are subdirectly irreducible. In both lattices, $0(A)$ and $1(A)$ form a critical edge.*

Proof. Obvious.

Lemma 3. *Let A and B be atomic Boolean lattices with A finite. If $K(A) \in \mathbf{HS}(K(B))$, then $K(A)$ is isomorphic to the sublattice of $K(B)$ generated by some subset of the form $\{a_i(A) \mid i \in I\} \cup \{d_i(A') \mid i \in I\}$ for some set I. In particular, $|A| \leqq |B|$.*

Proof. Let $K(A)$ be isomorphic under φ to S/Θ, where S is a sublattice of $K(B)$, and let Θ be a congruence of S. We claim that for any atom $a_i(A)$ of $K(A)$, $a_i(A)\varphi = \{a_j(B)\}$ for some atom $a_j(B)$ of $K(B)$. Indeed, let $a_i(A)\varphi = [x]\Theta$ for some x in S. Suppose that the claim fails. Since $a_i(A)$ is not the zero of $K(A)$, x can be chosen so that $x > a_j(B)$ for some $a_j(B)$ in $K(B)$ or $x \geqq 0(B')$. But $[x)$ is distributive in $K(B)$, which would imply that $[a_i(A))$ is distributive in $K(A)$, contradicting $|A| \geqq 8$, and verifying the claim. Similarly, for a dual atom $d_i(A')$ of $K(A)$, $d_i(A')\varphi = \{d_j(B')\}$ for some dual atom $d_j(B')$ of $K(B)$. The lemma now follows.

Lemma 4. *If A and B are atomic Boolean lattices with A finite, then $L(A) \notin \mathbf{HS}(K(B))$.*

Proof. Indeed, if $L(A) \in \mathbf{HS}(K(B))$, then $K(A) \in \mathbf{HS}(K(B))$. By Lemma 3, $K(A)$ is embedded into $K(B)$, with the unit of $K(A)$ going into the unit of $K(B)$. So there is no room for $u(A)$ and $v(A)$ in $K(B)$.

Lemma 5. *Let A and B be atomic Boolean lattices with A finite. If $L(A) \in \mathbf{HS}(L(B))$, then A and B are isomorphic.*

Proof. Let $L(A)$ be represented as S/Θ, where S is a sublattice of $L(B)$ and let Θ be a congruence relation of S. $u(B) \in S$, because otherwise S is a sublattice of $K(A)$, contradicting Lemma 4. Again, by Lemma 4, $u(B)$ cannot be congruent to an element of B' under Θ; nor can it be congruent to $v(B)$ because then the quotient could not contain $L(A)$. Thus $[u(B)]\Theta = \{u(B)\}$ represents $u(A)$; it follows, that $(S - \{u(B), v(B)\})/\Theta$ represents $K(A)$, hence $K(A) \in \mathbf{HS}(K(B))$. By Lemma 3, $K(A)$ is a specific type of sublattice of $K(B)$, the dual atoms $d_1(A'), d_2(A'), \ldots$ of A' correspond to dual atoms of B'. If A and B are not isomorphic, then A has fewer atoms, so their meet, $0(A')$ will not map onto $0(B')$, and will not be below $u(B)$, a contradiction.

Now we can state and prove the theorem of this section:

Theorem 4. $\mathbf{D}^2$ *has continuumly many subvarieties.*

Proof. Let N be a set of natural numbers $\geqq 3$. Let $\mathbf{V}(N)$ be the variety of lattices generated by the $L(A)$, where $|A|=2^n$ for some $n \in N$. We claim that for a finite Boolean lattice B, $L(B) \in \mathbf{V}(N)$ iff $|B|=2^m$ for some $m \in N$. This claim proves the theorem.

To verify the claim, let $L(B) \in \mathbf{V}(N)$. By Lemma 2, $L(B)$ is subdirectly irreducible, hence by Jónsson's Lemma, $L(B) \in \mathbf{HS}(L)$, where L is an ultraproduct of $L(A)$ with $|A|=2^n$, $n \in N$. However, the class of all $L(A)$, where A is an atomic Boolean lattice, is first-order definable. Hence, $L=L(A)$. $L(A) \in \mathbf{HS}(L(B))$ contradicts Lemma 5, unless A and B are isomorphic. This verifies the claim.

References

[1] R. Freese, Projective geometries as projective modular lattices, *Trans. Amer. Math. Soc.*, **251** (1979), 329—342.

[2] G. Grätzer, *General Lattice Theory*, Pure and Applied Mathematics series, Academic Press (New York, N. Y., 1978).

[3] G. Grätzer and D. Kelly, A survey of products of lattice varieties, in: *Contributions to lattice theory* (Proc. Conf. Szeged, 1980), Colloq. Math. Soc. J. Bolyai, vol. 33, North-Holland (Amsterdam, 1983); pp. 457—472.

[4] G. Grätzer and D. Kelly, Products of lattice varieties, *Algebra Universalis*, **21** (1986), 33—45.

[5] G. Grätzer and D. Kelly, The free $\mathfrak{m}$-lattice on the poset H, *Order*, **1** (1984), 47—65.

[6] T. A. Harrison, *P a nontrivial modular lattice variety and V a nonmodular lattice variety such that $V \circ P$ is a variety implies V is the variety of all lattices*, Ph. D. Thesis, University of Hawaii (1985).

[7] B. Jónsson and I. Rival, Lattice varieties covering the smallest non-modular variety, *Pacific J. Math.*, **82** (1979), 463—478.

[8] V. B. Lender, The groupoid of prevarieties of lattices, *Sibirsk. Mat. Zh.*, **16** (1975), 1214—1223. (Russian) (English translation: *Siberian Math. J.*, **16** (1975), 930—937.)

[9] R. McKenzie and D. Hobby, *The structure of finite algebras (tame congruence theory)*, Manuscript, 1985.

[10] S. V. Polin, On identities in congruence lattices of universal algebras, *Mat. Zametki*, **22** (1977), 443—451. (Russian)

[11] I. Rival and R. Wille, Lattices freely generated by partially ordered sets: which can be drawn? *J. Reine Angew. Math.*, **310** (1979), 55—80.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MANITOBA
WINNIPEG, MANITOBA, CANADA, R3T 2N2

Acta Sci., Math., 51 (1987), 81—91

Lattices whose congruence lattice is relative Stone

M. HAVIAR and T. KATRIŇÁK

Dedicated to the memory of A. Huhn

1. Introduction. A well-known fact about congruence lattices of lattices is that they are distributive, and hence, relatively pseudocomplemented. It is natural to seek for a characterization of those lattices whose congruence lattices satisfy some identities formulated in terms of (relative) pseudocomplements. For example, one can try to find those lattices whose congruence lattices are Boolean (cf. G. BIRKHOFF [2; Problem 39]), Stonean, relative Stonean, respectively. (For algebras see T. KATRIŇÁK and S. EL-ASSAR [17].)

There are three solutions of Birkhoff's problem: T. TANAKA, P. CRAWLEY (cf. [17]), G. GRÄTZER and E. T. SCHMIDT [6] have characterized those lattices whose congruence lattices are Boolean. Lattices whose congruence lattices form Stone lattices have been characterized in T. KATRIŇÁK [13]. In this note we answer a similar question: Characterize those lattices whose congruence lattices form relative Stone lattices.

The solution will be presented in terms of weak projectivity (Section 3). In Section 4 we first investigate lattices of the form $\mathbf{2}^P$. Then, as a consequence, we obtain an answer to the question formulated above for lattices L, with congruence lattices $\mathrm{Con}\,(L)$ isomorphic to some $\mathbf{2}^P$. In the last section we prove that a distributive lattice L has a relative Stone $\mathrm{Con}\,(L)$ if and only if $\mathrm{Con}\,(L)$ is Boolean.

2. Preliminaries. Let $\mathrm{Con}\,(L)$ denote the lattice of all congruence relations on a lattice L with Δ and ∇, the smallest and the largest congruence relation, respectively. It is well known (cf. [2] or [5]) that $\mathrm{Con}\,(L)$ satisfies the infinite distributivity

$$\theta \wedge \bigvee(\alpha_i\colon\ i\in I) = \bigvee(\theta\wedge\alpha_i\colon\ i\in I)$$

for any $\theta,\ \alpha_i\in\mathrm{Con}\,(L)$. It follows that for any $\alpha,\ \beta\in\mathrm{Con}\,(L)$ there exists a largest $\tau\in\mathrm{Con}\,(L)$ (the relative pseudocomplement) such that $\alpha\wedge\tau\leqq\beta$. Clearly, $\tau=$

Received June 30, 1986.

$=\bigvee(\delta\colon \alpha\wedge\delta\leqq\beta)$. (Notation: $\tau=\alpha_*\beta$.) Thus, Con (L) is a complete relatively pseudocomplemented lattice (or a complete Heyting algebra).

A bounded lattice $(L;\ \vee,\ \wedge, 0, 1)$, henceforth simply L, is called *pseudocomplemented* (=PCL) if L can be equipped with a unary operation * characterized by the property:

$$a\wedge x=0 \quad \text{if and only if} \quad x\leqq a^*.$$

A distributive PCL L is called a *Stone* lattice if

$$x^*\vee x^{**}=1$$

for every $x\in L$. Evidently, every Boolean lattice (algebra) is a Stone lattice, because a Boolean algebra is a distributive PCL satisfying the identities: $x=x^{**}$ and $x\vee x^*=1$. A lattice L is said to be *relative Stone* if every interval of L is a Stone lattice.

Every Heyting algebra is a PCL. In particular, the congruence lattice Con (L) of a lattice L is a (distributive) PCL, in which

$$\alpha^*=\alpha_*\Delta$$

for every $\alpha\in\mathrm{Con}\,(L)$.

We shall use the notation $a/b\to c/d$ for the weak projectivity of quotients (see [5]). All undefined terms as well as general lattice theoretic results may be found in G. Birkhoff [2] or G. Grätzer [5] or in E. T. Schmidt [18].

3. The general case. We begin with some definitions.

Definition. 1 ([16; Definition 2.1]). Let L be a lattice, $\pi\in\mathrm{Con}\,(L)$ and a/b, u/v quotients of L. Then L is said to be *π-almost weakly modular* whenever $a/b\to u/v$ and $u\not\equiv v(\pi)$ imply the existence of a subquotient $a_1/b_1\subseteqq a/b$ with $a_1\not\equiv b_1(\pi)$ such that for every quotient r/s with $r\not\equiv s(\pi)$ and $a_1/b_1\to r/s$ there exists a quotient z/t with $r/s\to z/t$, $u/v\to z/t$ and $z\not\equiv t(\pi)$. (See Figure 1.)

$$\begin{tikzcd}
a_1/b_1 \subseteqq a/b \arrow[r] \arrow[d] & u/v \arrow[d, dotted] \\
r/s \arrow[r, dotted] & z/t
\end{tikzcd}$$

Figure 1

"Almost weakly modular" will mean "Δ-almost weakly modular".

In [13; Definition 8] there is a slightly different definition of the notion of almost weak modularity: For any nontrivial quotients a/b, c/d, u/v of a lattice L satisfying $a/b\to u/v$, $c/d\to u/v$ there exists a nontrivial subquotient $a_1/b_1\subseteqq a/b$ such that for every $a_1/b_1\to r/s$, $r\neq s$, there exists a nontrivial quotient z/t with $r/s\to z/t$ and $c/d\to z/t$. (See Figure 2.)

$$\begin{tikzcd} a_1/b_1 \subseteq a/b \arrow[r] \arrow[d] & u/v & c/d \arrow[l] \arrow[d, dotted] \\ r/s \arrow[rr, dotted] & & z/t \end{tikzcd}$$

Figure 2

Lemma 1. *Every almost weakly modular lattice is almost weakly modular in the sense of* [13, Definition 8], *and vice versa.*

The proof is straightforward.

Definition 2. Let L be a lattice and $\theta, \pi \in \mathrm{Con}(L)$. Then θ is said to be *π-weakly separable* if $\pi \leqq \theta$ and for any $a < b$ in L there exists a chain $a = z_0 \leqq z_1 \leqq \leqq \ldots \leqq z_n = b$ such that for each i either

(i) $z_{i+1}/z_i \to u/v$ and $u \equiv v(\theta)$ imply $u \equiv v(\pi)$ or

(ii) for every subquotient $r/s \subseteq z_{i+1}/z_i$ with $r \not\equiv s(\pi)$ there exists a quotient u/v satisfying $r/s \to u/v$, $u \equiv v(\theta)$ and $u \not\equiv v(\pi)$.

"Weakly separable" means "Δ-weakly separable".

It is easy to verify that every weakly modular lattice is almost weakly modular (cf. [5], [6]). Similarly, every separable congruence relation of a weakly modular lattice is almost separable (cf. [6] and [16]).

Using Lemma 1 we can reformulate a result from [13].

Theorem 1 ([13; Theorem 4]). *Let L be a lattice. Then* Con (L) *is a Stone lattice if and only if*

(i) *L is almost weakly modular and*

(ii) *every congruence relation of L is weakly separable.*

Lemma 2. *Let L be a lattice and $\pi \in \mathrm{Con}(L)$. Then L is π-almost weakly modular if and only if the factor lattice L/π is almost weakly modular.*

Proof. Suppose that L is π-almost weakly modular. Take $\bar{a}/\bar{b} \to \bar{u}/\bar{v}$ and $\bar{u} \neq \bar{v}$ in L/π. Then there exist $a \in \bar{a}$, $b \in \bar{b}$, $u \in \bar{u}$ and $v \in \bar{v}$ such that $a > b$, $u > v$ and $a/b \to u/v$, $u \not\equiv v(\pi)$ in L. Moreover, there exists a subquotient $a_1/b_1 \subseteq a/b$ with $a_1 \not\equiv b_1(\pi)$ having the properties described in Definition 1. Clearly $\bar{a}_1/\bar{b}_1$ is a subquotient of $\bar{a}/\bar{b}$ if $\bar{a}_1 = [a_1]\pi$ and $\bar{b}_1 = [b_1]\pi$. Now, as $\bar{a}_1/\bar{b}_1 \to \bar{r}/\bar{s}$ in L/π, there exists a quotient r/s in L with $r \in \bar{r}$, $s \in \bar{s}$, $a_1/b_1 \to r/s$ and $r \not\equiv s(\pi)$. Eventually, there exists a quotient z/t in L such that $z \not\equiv t(\pi)$, $r/s \to z/t$ and $u/v \to z/t$. Clearly, $\bar{r}/\bar{s} \to \bar{z}/\bar{t}$ and $\bar{u}/\bar{v} \to \bar{z}/\bar{t}$ in L/π.

The converse statement can be established in the same manner.

The following lemma can be verified in the same way as Lemma 2.

Lemma 3. *Let L be a lattice and $\pi, \theta \in \mathrm{Con}(L)$. Let $\pi \leqq \theta$. Then θ is π-weakly separable if and only if θ/π is a weakly separable congruence relation of L/π.*

Lemma 4. *Let L be a bounded distributive lattice. Then L is a relative Stone lattice if and only if for every $a \in L$ ($L=[0, 1]$) the interval $[a, 1]$ is a Stone lattice.*

Proof. The statement follows from the following observation: $[0, a]$ is a Stone lattice for every $a \in L$ whenever L is a Stone lattice. Really, $x^+ = x^* \wedge a$ is the pseudocomplement of $x \in [0, a]$ in $[0, a]$. Therefore, $x^{++} = (x^* \wedge a)^* \wedge a = x^{**} \wedge a$. Now, $x^+ \vee x^{++} = (x^* \vee x^{**}) \wedge a = a$, and the proof is complete.

Theorem 2. *Let L be a lattice. Then* $\mathrm{Con}(L)$ *is a relative Stone lattice if and only if for every $\pi \in \mathrm{Con}(L)$*

(i) *L is π-almost weakly modular and*

(ii) *every congruence relation θ of L is π-weakly separable.*

Proof. Owing to Lemma 4, $\mathrm{Con}(L)$ is a relative Stone lattice if and only if $\mathrm{Con}(L/\pi)$ is a Stone lattice for every $\pi \in \mathrm{Con}(L)$. The rest of the proof follows from Theorem 1 and Lemmas 2, 3.

π-almost weak modularity is a rather complicated condition. It can be somewhat simplified for semi-discrete lattices. This will be done in the next section.

4. Congruence lattices of the form $\mathbf{2}^P$. We shall start with some results on lattices of the form $\mathbf{2}^P$. There are several characterizations of $\mathbf{2}^P$ (see [2] and [8]).

Let P be a poset. $\mathbf{2}^P$ denotes the lattice of all isotone functions defined on P with values in the chain $\mathbf{2}$ of two elements, where $(f \vee g)(x) = f(x) \vee g(x)$, $(f \wedge g)(x) = f(x) \wedge g(x)$ for any $f, g \in \mathbf{2}^P$ and every $x \in P$.

Again, if P is a poset, then a subset Q of P is said to be *decreasing (increasing)* if $x \in Q$, $y \leqq x$ in P ($x \in Q$, $y \geqq x$ in P) imply $y \in Q$. $d(P)$ will denote the set of all decreasing subsets of P. $d(P)$ is a complete lattice in which the complete join and meet coincide with the set-theoretical join and meet. Dually, $i(P)$ will denote the set of all increasing subsets of P, which is a complete lattice with respect to the set-theoretical join and meet.

$\bar{P}$ denotes the dual poset of P. If $U \subseteq P$ then $[U) = \{x \in P\colon x \geqq y \text{ for some } y \in U\}$. Dually we define $(U]$. Clearly, $[U) \in i(P)$, $(U] \in d(P)$, $P-(U] \in i(P)$ and $P-[U) \in d(P)$. (Here "$-$" denotes the set-theoretical difference. $\emptyset$ is the void set.)

We get immediately from the definitions:

Lemma 5. *Let P be a poset. Then*

(i) $d(P) \cong \overline{i(P)} \cong i(\bar{P})$;

(ii) $\mathbf{2}^P \cong i(P)$.

Theorem 3. *Let P be a poset. Then $i(P)$ is a double Heyting algebra, i.e., $i(P)$ is relatively pseudocomplemented and dually relatively pseudocomplemented. More precisely, for $U, V \in i(P)$ we have:*

(i) $U_* V = P-(U-V]$ (see also [14; 2.1]);

(ii) $U^* = U_* \emptyset = P-(U]$;

(iii) $U_+ V = [V-U)$ (= the dual relative pseudocomplement).

Proof. The first part of the statement follows from the fact that $i(P)$ is isomorphic to a complete sublattice of an atomic complete Boolean algebra (see [8; Theorem 3]). We shall prove (i). Evidently,

$$U \cap (P-(U-V]) \subseteqq U \cap V \subseteqq V.$$

Conversely, assume $U \cap W \subseteqq V$ for some $W \in i(P)$. Suppose to the contrary that $(U-V] \cap W \neq \emptyset$. Then there exists $t \in (U-V] \cap W$. Hence, there exists $x \geqq t$ with $x \in U-V$. Clearly $x \in W$. Therefore,

$$x \in (U-V) \cap W \subseteqq U \cap W \subseteqq V,$$

a contradiction. Thus, $W \subseteqq P-(U-V]$ and (i) is established.

(ii) follows immediately from (i). (iii) can be proven in a similar way as (i).

Corollary 1. *Let P be a poset, $U, V \in d(P)$. Then $U_* V = P-[U-V)$ and $U_+ V = (V-U]$.*

Corollary 2. *Let $f, g \in \mathbf{2}^P$. Then*

(i) $(f_* g)(x) = 0$ *if and only if there exists* $y \geqq x$ *such that* $f(y)=1$ *and* $g(y)=0$;

(ii) $(f_+ g)(x) = 1$ *if and only if there exists* $y \leqq x$ *such that* $f(y)=0$ *and* $g(y)=1$.

How can the lattices $\mathbf{2}^P$ be related to congruence lattices? This can be found in [6] and [8]:

Let Q be the set of all prime quotients of the lattice L. The elements of Q are denoted by p, q, r. If $p=a/b$, $q=c/d$ are prime quotients and $a/b \to c/d$, then we write $p \to q$. The elements of Q under the relation $\to$ are quasiordered. Define the relation $\sim$ by $p \sim q$ if and only if $p \to q$ and $q \to p$. Then $P = Q/\sim$ is partially ordered.

If in L all bounded chains are finite then we speak of a *discrete* lattice. Further, if in L between all comparable pairs of elements there exists a finite maximal chain, then we call L *semi-discrete*.

Lemma 6 ([6; Lemma 19, Theorem 13]). *For any lattice L, there exists a poset P such that $\mathbf{2}^P$ is a complete homomorphic image of* Con(L). *If L is semi-discrete then* Con$(L) \cong \mathbf{2}^P$.

Lemma 7 ([8; Theorem 2]). *Let P be a poset. Then there exists a section complemented locally finite lattice L such that* Con$(L) \cong \mathbf{2}^P$.

Now we can present the first result.

Theorem 4. *Let P be a poset. Then $\mathbf{2}^P$ is a Stone lattice if and only if $u \leqq a$ and $u \leqq b$ in P imply the existence of an element $s \in P$ such that $a \leqq s$ and $b \leqq s$.*

Proof. Assume that $\mathbf{2}^P$ is a Stone lattice. Therefore, $i(P)$ is a Stone lattice (Lemma 5). Take elements $u, a, b \in P$ with $u \leqq a$ and $u \leqq b$. Suppose to the contrary that there is no $s \in P$ such that $a \leqq s$ and $b \leqq s$. Take $[a) \in i(P)$. Therefore,

$$[a)^* = P-([a)] \quad \text{and} \quad [a)^{**} = P-([a)^*]$$

by Theorem 3. Since $i(P)$ is a Stone lattice, we have

$$[a)^* \cup [a)^{**} = P.$$

On the other hand, by the hypothesis, $u \in [a)^*$ and $b \in [a)^*$. Since $b \in [a)^*$, we get $u \notin [a)^{**}$, a contradiction. Thus $[a) \cap [b) \neq \emptyset$.

Conversely, suppose that the stated condition holds. Consider $U \in i(P)$. Therefore,

$$U^* \cup U^{**} = (P-(U]) \cup (P-(U^*]).$$

Suppose to the contrary that $U^* \cup U^{**} \neq P$. Then there exists $u \in P-(U^* \cup U^{**})$. Therefore, $u \in (U] \cap (U^*]$. There exist $a \in U$ and $b \in U^*$ such that $u \leqq a$ and $u \leqq b$. By the hypothesis

$$a \leqq s \quad \text{and} \quad b \leqq s \quad \text{for some} \quad s \in P.$$

Therefore, $s \in U \cap U^* = \emptyset$, a contradiction. Thus, $U^* \cup U^{**} = P$ and $i(P)$ is a Stone lattice.

Corollary ([13; Corollary to Theorem 4]). *Let L be a semi-discrete lattice. Then* Con (L) *is a Stone lattice if and only if for any prime quotients p, q, r of L satisfying $p \to q$ and $p \to r$ there exists a prime quotient s of L such that $q \to s$ and $r \to s$.* (See Figure 3.)

$$\begin{tikzcd} p \arrow[r] \arrow[d] & r \arrow[d, dotted] \\ q \arrow[r, dotted] & s \end{tikzcd}$$

Figure 3

Theorem 5. *Let P be a poset. Then $\mathbf{2}^P$ is a relative Stone lattice if and only if $u \leqq a$ and $u \leqq b$ in P imply $a \leqq b$ or $b \leqq a$, i.e., for every $u \in P$, $[u)$ is a chain.*

Proof. Assume that $\mathbf{2}^P$ is a relative Stone lattice. This means by Lemma 5 that $i(P)$ is a relative Stone lattice. Take elements $u, a, b \in P$ with $u \leqq a$ and $u \leqq b$.

Consider $[a), [b) \in i(P)$ in the interval $[[a) \cap [b), P]$ of $i(P)$. Let $[a)^+$ denote the pseudocomplement of $[a)$ in this interval. Evidently,

$$[a) \cap [b) \subseteqq [b) \subseteqq [a)^+, \quad [a) \cap [b) \subseteqq [a) \subseteqq [a)^{++}.$$

By the hypothesis $[a)^+ \cup [a)^{++} = P$. Suppose to the contrary that the elements a and b are incomparable. Therefore,

$$[a) \neq [a) \cap [b) \neq [b).$$

Clearly, $u \in [a)^+$ or $u \in [a)^{++}$. But $u \in [a)^+$ implies $a \in [a)^+$, as $u \leqq a$. Hence,

$$a \in [a) \cap [a)^+ = [a) \cap [b) \neq [a),$$

a contradiction. Therefore, $u \in [a)^{++}$. It follows that $b \in [a)^{++}$, as $u \leqq b$. Hence,

$$b \in [a)^+ \cap [a)^{++} = [a) \cap [b) \neq [b),$$

a contradiction. Thus, the elements a and b are comparable.

Conversely, let $[u)$ be a chain for every $u \in P$. According to Lemma 4 it is enough to show that for every $U \in i(P)$ the interval $[U, P]$ forms a Stone lattice. Take $V \in [U, P]$. Consider the pseudocomplements V^+ and V^{++} of V and V^+ in $[U, P]$, respectively. It is easy to verify that

$$V^+ = V_* U \quad \text{and} \quad V^{++} = V^+{}_* U = (V_* U)_* U.$$

Assume to the contrary that $V^+ \cup V^{++} \neq P$. Then there exists $u \in P$ such that $u \notin V^+$ and $u \notin V^{++}$. This implies

$$u \in (V - U] \quad \text{and} \quad u \in (V^+ - U],$$

by Theorem 3. There exist $a \in V - U$ and $b \in V^+ - U$ such that $u \leqq a$ and $u \leqq b$. By the hypothesis $a \leqq b$ or $b \leqq a$. Now, $a \leqq b$ yields

$$b \in V \cap V^+ = U,$$

which contradicts $b \notin U$. The remaining case $b \leqq a$ implies again $a \in V \cap V^+ = U$, which is impossible. Thus $V^+ \cup V^{++} = P$, and $i(P)$ is a relative Stone lattice.

Corollary. *Let L be a semi-discrete lattice. Then* Con (L) *is a relative Stone lattice if and only if for any prime quotients p, q of L satisfying $p \to q$ and $p \to r$ either $q \to r$ or $r \to q$ holds.* (See Figure 4.)

$$\begin{tikzcd} p \arrow[r] \arrow[d] & q \arrow[dl, dotted, bend left] \\ r \arrow[ur, dotted, bend left] & \end{tikzcd}$$

Figure 4

A bounded relative Stone lattice L can be considered as a Heyting algebra $(L;\vee,\wedge,{}_*,0,1)$. More precisely,

Lemma 8 ([15; 2.9 and 2.10]). *Let L be a bounded distributive lattice. Then L is a relative Stone lattice if and only if L is a Heyting algebra satisfying the identity*

$$x_*y\vee y_*x=1.$$

Since the class of all Heyting algebras is equational, we see that the class of all bounded relative Stone lattices forms a variety. In [10] it has been shown that the lattice of all subvarieties of the variety of bounded relative Stone lattices is isomorphic to the chain of type $\omega+1$. In addition, a Heyting algebra L belongs to the n-th $(n\geqq 2)$ subvariety of the variety of all relative Stone lattices if and only if it satisfies the identity

$$(E_n)\qquad (x_{1*}x_2)\vee(x_{2*}x_3)\vee\ldots\vee(x_{n*}x_{n+1})=1.$$

Note that the subvariety satisfying (E_2) is exactly the class of all Boolean algebras.

Theorem 6. *Let P be a poset. Then $\mathbf{2}^P$ satisfies the identity (E_n) for $n\geqq 2$ if and only if*

(i) *$\mathbf{2}^P$ is a relative Stone lattice and*

(ii) *any chain of P possesses at most $n-1$ elements.*

Proof. Assume that $\mathbf{2}^P$ satisfies the identity (E_n). Set $x=x_1=x_3=\ldots$ and $y=y_2=\ldots$ in (E_n). It follows that $x_*y\vee y_*x=1$ is true, and (i) is established (Lemma 8). Now, suppose to the contrary that P contains an n-element chain

$$x_1<x_2<\ldots<x_n.$$

In $i(P)$ we get the following $(n+1)$-element chain

$$[x_1)\supset[x_2)\supset\ldots\supset[x_n)\supset\emptyset.$$

Using Theorem 3 we obtain

$$[x_i)_*[x_{i+1})=P-([x_i)-[x_{i+1})],\quad [x_n)_*\emptyset=P-([x_n)].$$

Clearly, $x_1\in([x_i)-[x_{i+1})]$ and $x_1\in([x_n)]$ for every $i=1,\ldots,n-1$. Therefore,

$$x_1\notin[x_1)_*[x_2)\cup[x_2)_*[x_3)\cup\ldots\cup[x_n)_*\emptyset=P,$$

which is impossible. Thus, (i) is true.

Conversely, assume (i) and (ii). Take $U_1,\ldots,U_{n+1}\in i(P)$. We shall investigate $W=U_{1*}U_2\cup\ldots\cup U_{n*}U_{n+1}$ in $i(P)$. By Theorem 3,

$$W=(P-(U_1-U_2])\cup\ldots\cup(P-(U_n-U_{n+1}])=P-((U_1-U_2]\cap\ldots\cap(U_n-U_{n+1}]).$$

Assume to the contrary that $W \neq P$. Then there exists

$$a \in (U_1 - U_2] \cap (U_2 - U_3] \cap \ldots \cap (U_n - U_{n+1}].$$

It follows that there exist $x_i \in U_i - U_{i+1}$ with $a \leqq x_i$ for all $i = 1, \ldots, n$. By (i) and Theorem 5 $\{a, x_1, \ldots, x_n\}$ is a chain. We claim that

$$a \leqq x_1 < x_2 < \ldots < x_n.$$

Indeed, $x_i \in U_i - U_{i+1}$ means $x_i \in U_i$ and $x_i \notin U_{i+1}$. Since x_i, x_{i+1} are comparable elements, $x_i \notin U_{i+1}$, $x_{i+1} \in U_{i+1}$ and U_{i+1} is increasing, we see that $x_i < x_{i+1}$ for every $i = 1, \ldots, n$, as claimed. But $x_1 < \ldots < x_n$ is an n-element chain, which contradicts (ii). Thus $i(P)$ satisfies (E_n) and the proof is complete.

Corollary 1 ([6], [8]). *$\mathbf{2}^P$ is a Boolean algebra if and only if P is unordered.*

Proof. $\mathbf{2}^P$ is Boolean if and only if $\mathbf{2}^P$ satisfies (E_2).

Corollary 2 ([6], [11]). *Let L be a semi-discrete lattice. Then* Con (L) *is a Boolean algebra if and only if for any prime quotients p, q of L, $p \to q$ implies $q \to p$.*

Corollary 3. *Let L be a semi-discrete lattice. Then* Con (L) *satisfies the identity (E_n) for $n \geqq 2$ if and only if* Con (L) *is a relative Stone lattice and for any prime quotients $p, q_1, \ldots, q_{n-1}$ of L satisfying*

$$p \to q_i \quad \textit{for all} \quad i = 1, \ldots, n-1$$

either $q_i \to p$ or $q_i \to q_j$ and $q_j \to q_i$ $(i \neq j)$ holds for some $i, j \in \{1, \ldots, n-1\}$.

Remark. Lemma 7 and Theorems 4—6 enable us to construct lattices L with Con (L) a Stone lattice, a relative Stone lattice or a lattice satisfying (E_n) for some $n \geqq 2$, respectively.

5. Congruence lattices of distributive lattices. We shall need the following two classical results.

Lemma 9 ([7], [9]). *To any distributive lattice L there exists a generalized Boolean algebra B having the properties:*

(i) *L is a sublattice of B,*

(ii) Con $(L) \cong$ Con (B),

(iii) *if the interval $[a, b]$ of L is of finite length, then $[a, b]$ has the same length as an interval of B.*

Lemma 10 ([9], [6]). *Let L be a distributive lattice. Then* Con (L) *is a Boolean algebra if and only if L is discrete.*

The proof of the following statement is straightforward.

Lemma 11. *Let H be a Heyting algebra, $b\in H$. Then*
(i) $((x\wedge b)_*(y\wedge b))\wedge b=(x_*y)\wedge b$,
(ii) $((b];\ \vee,\ \wedge,\ \to, 0, b)$ *is a Heyting algebra if* $x\to y=(x_*y)\wedge b$,
(iii) *the map* $\varphi\colon x\mapsto x\wedge b$ *is an epimorphism between the Heyting algebras* H *and* $(b]$.

Theorem 7. *Let L be a distributive lattice. Then* $\mathrm{Con}\,(L)$ *is a relative Stone lattice if and only if* $\mathrm{Con}\,(L)$ *is a Boolean algebra.*

Proof. Suppose that $\mathrm{Con}\,(L)$ is a relative Stone lattice. L finite yields that $\mathrm{Con}\,(L)$ is a Boolean algebra (Lemma 10). We can assume L infinite. We shall first investigate the case that L is a bounded lattice. Then by [7; Corollary 2 to Theorem 2] the lattice B from Lemma 9 is a Boolean algebra. Since $\mathrm{Con}\,(L)\cong\mathrm{Con}\,(B)$, we conclude that every homomorphic image of the Boolean algebra B is complete (see Lemma 4 and [4; Theorem 4] or [12; Theorem 6]; or [1; Corollary 4 to Theorem 2]). But this contradicts the following statement of Ph. Dwinger [3]: *Every infinite complete Boolean algebra has an incomplete homomorphic image.* Thus, L infinite and bounded cannot occur.

Now, suppose that L is infinite having no largest (smallest) element. Assume to the contrary that L is not discrete. There exists an interval $[a, b]$ of L with an infinite chain. Consider the generalized Boolean algebra B satisfying $\mathrm{Con}\,(L)\cong\mathrm{Con}\,(B)$ from Lemma 9. Denote by B_1 the interval $[0, b]=(b]$ of B. Clearly, B_1 is an infinite Boolean algebra. Every congruence relation of B (B_1) is uniquely determined by its kernel. Therefore,

$$\mathrm{Con}\,(B)\cong I(B) \quad\text{and}\quad \mathrm{Con}\,(B_1)\cong I(B_1),$$

where $I(B)$ denotes the lattice of all ideals of B. Since $B_1\in I(B)$, we see that $I(B_1)$ is isomorphic to the interval $[(0], B_1]$ of $I(B)$. Now we can apply Lemma 11: $I(B_1)$ is an epimorphic image of $I(B)$. $I(B)$ satisfies the identity

$$x_*y\vee y_*x=1$$

(Lemma 8). Hence, $I(B_1)$ also satisfies this identity. It follows again by Lemma 8 that $\mathrm{Con}\,(B_1)$ is a relative Stone lattice. Therefore, B_1 is an infinite Boolean algebra such that every homomorphic image of B_1 is complete. But this is not true by the mentioned theorem of Ph. Dwinger [3]. Thus, L is discrete. By Lemma 10 $\mathrm{Con}\,(L)$ is a Boolean algebra.

The converse statement is trivial.

References

[1] R. Beazer, Lattices whose ideal lattice is Stone, *Proc. Edinburgh Math. Soc.*, **26** (1983), 107—112.

[2] G. Birkhoff, *Lattice Theory*, Third Ed., Colloq. Publ., vol. 25, Amer. Math. Soc. (Providence, R. I., 1967).

[3] Ph. Dwinger, On the completeness of the quotient algebras of a complete Boolean algebra. I, *Indag. Math.*, **20** (1958), 448—456.

[4] O. Frink, Pseudo-complements in semi-lattices, *Duke Math. J.*, **29** (1962), 505—514.

[5] G. Grätzer, *General Lattice Theory*, Birkhäuser Verlag (Basel, 1978).

[6] G. Grätzer and E. T. Schmidt, Ideals and congruence relations in lattices, *Acta Math. Acad. Sci. Hungar.*, **9** (1958), 137—175.

[7] G. Grätzer and E. T. Schmidt, On the generalized Boolean algebra generated by a distributive lattice, *Indag. Math.*, **20** (1958), 547—553.

[8] G. Grätzer and E. T. Schmidt, On congruence lattices of lattices, *Acta Math. Acad. Sci. Hungar.*, **13** (1962), 179—185.

[9] J. Hashimoto, Ideal theory for lattices, *Math. Japon.*, **2** (1952), 149—186.

[10] T. Hecht and T. Katriňák, Equational classes of relative Stone algebras, *Notre Dame J. Formal Logic*, **13** (1972), 248—254.

[11] J. Jakubík, Relácie kongruentnosti a slabá projektívnosť vo zväzoch, *Časop. pěst. mat.*, **80** (1955), 206—216. (Slovak)

[12] T. Katriňák, Notes on Stone lattices. I, *Mat. fyz. časop.*, **16** (1966), 128—142. (Russian)

[13] T. Katriňák, Notes on Stone lattices. II, *Mat. časop.*, **17** (1967), 20—37. (Russian)

[14] T. Katriňák, *M*-Polaren in halbgeordneten Mengen, *Časop. pěst. mat.*, **95** (1970), 416—419.

[15] T. Katriňák, Die Kennzeichnung der distributiven pseudokomplementären Halbverbände, *J. reine angew. Math.*, **241** (1970), 160—179.

[16] T. Katriňák, Eine Charakterisierung der fast schwach modularen Verbände, *Math. Z.*, **114** (1970), 49—58.

[17] T. Katriňák and S. El-Assar, Algebras with Boolean and Stonean congruence lattices, *Acta Math. Hungar.*, **48** (1986), 301—316.

[18] E. T. Schmidt, *A Survey on Congruence Lattice Representations*, Teuber-Texte zur Math., Band 42 (Leipzig, 1982).

KATČ MFFUK
84 215 BRATISLAVA, CZECHOSLOVAKIA

Acta Sci. Math., 51 (1987), 93—101

Frames of permuting equivalences

CHRISTIAN HERRMANN

Dedicated to the memory of András Huhn

András Huhn established frames as the fundamental tool in the equational theory of modular lattices. In the present note we use this algebraic point of view for an easy approach to von Neumann's Coordinatization Theorem [11] — completing a program of FRINK [5] and JÓNSSON [9] based on the abelian group representation given by the Embedding Theorem.

This approach can be extended to permuting equivalence representations of lattices with spanning frames of order $n \geqq 3$. The loop associated with the net provides a module representation for the sublattice generated by the frame and its coordinate ring. From this we can derive a lattice identity separating lattices of permuting equivalences on finite sets and finite lattices having an (infinite) permuting equivalence representation.

I have to thank Ralph Freese and the referee for doing an excellent job.

1. The associated group

A *frame* Φ of order n in a lattice L consists of elements a_i, c_{ij} $(1 \leqq i \neq j \leqq n)$ such that

$$a_i \cap \sum_{k \neq i} a_k = \bigcap_k a_k = a_i \cap c_{ij}, \quad a_i + c_{ij} = a_i + a_j,$$

$$c_{ik} = c_{ki} = (c_{ij} + c_{jk}) \cap (a_i + a_k).$$

It is *spanning* if $\bigcap a_k$ and $\sum a_k$ are the bounds of L. For every module A we have the *canonical* frame in the lattice $L(A^n)$ of all submodules of A^n given by

$$\{(0, \ldots, \underset{i}{x}, \ldots, 0) \mid x \in A\}, \quad \{(0, \ldots, \underset{i}{x}, \ldots 0, \ldots, -\underset{j}{x}, \ldots, 0) \mid x \in A\}.$$

Received July 29, 1986, and in revised form October 21, 1986.

The *coordinate domain* R_{ij} of Φ in L consists of all r in L such that

$$r+a_j = a_i+a_j, \quad r\cap a_j = a_i\cap a_j,$$
$$r = (r+a_k)\cap(a_i+a_j) = (r+c_{ik})\cap(a_i+a_j) = (r+c_{jk})\cap(a_i+a_j)$$

for all $k\neq i,j$. For modular L the last identity is superfluous.

A frame Φ contained in the lattice $\Pi(E)$ of all partitions on the set E is called *permuting*, if any two of the a_i's permute and c_{ij} with a_i for all i, j (it suffices to consider a spanning tree of pairs $\{i,j\}$ for which c_{ij} and c_{jk} permute, too). Then, r is in R_{ij} if it permutes with all a_k, c_{ik}, and c_{jk} $(k\neq i,j)$ and is a complement of a_j in $[a_ia_j,\ a_i+a_j]$.

Often, we prefer to think of equivalence relations. In particular, with any subgroup B of an abelian group A we associate the congruence relation β on A given by $x\beta y$ iff $x-y\in B$. A closer look at the loop associated with a net yields

Theorem 1. *Let Φ be a permuting spanning frame of order $n\geqq 3$ in $\Pi(E)$. Then there is an abelian group A and a bijection $\varphi\colon E\to A^n$ mapping Φ onto congruences associated with the canonical frame of A^n and all coordinate domain elements onto congruences of A^n.*

Corollary 2. *A permuting frame of order $n\geqq 3$ in a partition lattice generates, together with all its coordinate domains, a complete sublattice of permuting equivalences.*

Proof. Denote the frame by α_i and ε_{ij}. In view of the permutability and independence of the α_i we may assume $E=A_1\times\ldots\times A_n$ with $(a_1, \ldots, a_n)\ \alpha_i\ (b_1, \ldots, b_n)$ iff $a_j=b_j$ for all $j\neq i$. Choose an element 0_i in A_i for each i. Then

$$f_{ij}(x) = y \quad \text{iff} \quad (0, \ldots, \underset{i}{x}, \ldots, 0)\ \varepsilon_{ij}\ (0, \ldots, \underset{j}{y}, \ldots, 0)$$

defines a bijection of A_i onto A_j mapping 0_i onto 0_j — due to the permutability of ε_{ij} with α_i and α_j. The normalization condition for the frame yields $f_{jk}\circ f_{ij}=f_{ik}$. Thus, we may identify A_i with A_j via f_{ij} to obtain $E=A^n$ with

$$(a_1, \ldots, a_n)\ \varepsilon_{ij}\ (b_1, \ldots, b_n) \quad \text{iff} \quad a_i = b_j \quad \text{and} \quad a_k = b_k \quad \text{for all} \quad k\neq i,j$$

provided that $a_j=0=b_i$. Namely, let $i=1$, $j=2$. Since $\varepsilon_{12}\subseteq\alpha_1+\alpha_2$ we may assume $a_k=b_k$ for all $k\geqq 3$. If these are 0 the claim is obvious. The general case reduces to this one since

$$a\,\varepsilon_{12}\,b \quad \text{iff} \quad (a_1, 0, \ldots, 0) \quad \varepsilon_{12}+\sum_{k\geqq 3}\alpha_k \quad (0, b_2, 0, \ldots, 0)$$

in view of $(\alpha_1+\alpha_2)\cap(\varepsilon_{12}+\sum_{k\geqq 3}\alpha_k)=\varepsilon_{12}$.

The 3-net (cf. Dénes and Keedwell [3]) $\alpha_i, \alpha_j, \varepsilon_{ij}$ on the $\alpha_i+\alpha_j$-class of $(0, \ldots, 0)$ yields for every $i\neq j$ and a, b in A a uniquely determined $c=a+_{ij}b$ in A such that

$$(0, \ldots, \underset{i}{a}, \ldots, 0, \ldots, \underset{j}{b}, \ldots, 0)\ \varepsilon_{ij}\ (0, \ldots, \underset{i}{c}, \ldots, 0).$$

Claim. A with 0 and $+_{ij}$ is an abelian group not depending on i, j.

We may assume $n=3$, $i \neq j \neq k \neq i$, e.g. $i=1$, $j=2$, $k=3$. Observe that

$$\alpha_i \cap (\varepsilon_{ij} + \varepsilon_{jk} + \varepsilon_{ik}) = \mathrm{id}$$

due to the frame relations. We have

$$a +_{ij} b = b +_{ji} a$$

since $(a +_{12} b, 0, 0)\, \varepsilon_{12}\, (a, b, 0)\, \varepsilon_{21}\, (0, b +_{21} a, 0)\, \varepsilon_{12}\, (b +_{21} a, 0, 0)$, and

$$a +_{ij} b = a +_{ik} b$$

since $(a +_{12} b, 0, 0)\, \varepsilon_{12}\, (a, b, 0)\, \varepsilon_{23}\, (a, 0, b)\, \varepsilon_{13}\, (a +_{13} b, 0, 0)$. Now, indepedence and commutativity follow. 0 is neutral, obviously. Due to the permutability of α_2 and ε_{12} for any a there is (b, c, d) with $(0, 0, 0)\ \varepsilon_{12}\ (b, c, d)\ \alpha_2\ (a, 0, 0)$, whence $b=a$, $d=0$, and $a+c=0$. Finally, associativity follows from

$$\big((a+b)+c, 0, 0\big)\, \varepsilon_{13}\, (a+b, 0, c)\, \varepsilon_{12}\, (a, b, c)\, \varepsilon_{23}\, (a, 0, b+c)\, \varepsilon_{13}\, \big(a+(b+c), 0, 0\big).$$

For ϱ in R_{12} we have $\varrho \subseteq \alpha_1 + \alpha_2$ whence

$$(x, y, c)\ \varrho\ (a, b, c) \quad \text{iff} \quad (x, y, d)\ \varrho\ (a, b, d)$$

since $(x, y, c)\ \varrho\ (a, b, c)$ implies $(x, y, d)\ \alpha_3\ (x, y, c)\ \varrho\ (a, b, c)\ \alpha_3\ (a, b, d)$ and since $\varrho = (\varrho + \alpha_3) \cap (\alpha_1 + \alpha_2)$. Also,

$$(x, y, 0)\ \varrho\ (a, b, 0) \quad \text{implies} \quad (x+u, y, 0)\ \varrho\ (a+u, b, 0)$$

since $(x+u, y, 0)\ \ \varepsilon_{13}\ (x, y, u)\ \ \varrho\ \ (a, b, u)\ \ \varepsilon_{13}\ (a+u, b, 0)$. With the corresponding property for the second component we get that $(x, y, z)\ \varrho\ (a, b, c)$ implies $(x+u, y+v, z+w)\ \varrho\ (a+u, b+v, c+w)$ which means that ϱ is a congruence of the group A^n.

2. The associated module

When studying faithful permutable representations of a lattice with a spanning frame of order $n \geqq 3$ we may, in view of Theorem 1, restrict attention to sublattices of $L(A^n)$ with canonical Φ, A an abelian group. Then, each element of the coordinate domain R_{ij} can be considered the graph of an endomorphism $x \mapsto -xr$ of A

$$\{(0, \ldots, \underset{i}{x}, \ldots, 0, \ldots, -\underset{j}{x}r, \ldots, 0) \mid x \in A\} =: r_{ij},$$

$r_{ik} = (r_{ij} + r_{jk}) \cap (a_i + a_k)$. With von Neumann's addition and multiplication

$$(s-r)_{ij} = \big((s_{ij} + c_{jk}) \cap (a_j + r_{ik}) + a_k\big) \cap (a_i + a_j),$$

$$(sr)_{ik} = (r_{ij} + s_{jk}) \cap (a_i + a_k),$$

the R_{ij} are isomorphic to a subring of the endomorphism ring of A. Also, if L is generated by Φ and the R_{ij} then it is contained in the lattice $L(A^n_T)$ of right T-submodules, where T is the commuting ring of R and A, i.e. the endomorphism ring of A_R. Finally, if S is a subring of T and A_S a cyclic module then identification of A_S with S yields a representation of L in the lattice $L({}_SS^n)$ of left S-submodules of S^n.

Proposition 3. *Let R be a commutative completely primary uniserial ring, Φ the canonical frame of $L=L({}_RR^n)$, and $\varphi\colon L\to\Pi(E)$ a representation with permuting and spanning frame $\varphi(\Phi)$. Then $|E|\leqq 1$ or there is a bijection ψ of E onto a direct sum ${}_RM$ of modules ${}_RR^n$ such that $\psi\circ\varphi$ becomes the diagonal embedding of L into $L({}_RM)$.*

These lattices L are particular primary lattices of type (n) in the sense of Jónsson and Monk [10]. The proposition allows to view every permuting equivalence representation as derived from the Jónsson—Monk coordinatization.

Proof. Since L is simple and generated by $\Phi\cup R_{12}$ (cf. [6], 2.9) it suffices to consider φL as a sublattice of $L(A^n_R)$ and $\varphi\Phi$ canonical. Now, A_R is a direct sum of cyclic modules which yields a decomposition of $\varphi\Phi$ (see Herrmann and Huhn [7], Section 2) and of the coordinate domain φR_{12}. Thus, all of φL is decomposed which means that we have a direct sum of representations with cyclic A_R's. The latter are full lattices $L({}_RR^n)$ since these are generated by $\Phi\cup R_{12}$.

By the *indices* of a partition β in α we mean the numbers of β-classes in the classes of α. If α and β permute then $\alpha\cap\beta\subseteq\beta$ and $\alpha\subseteq\alpha+\beta$ have the same set of indices.

Corollary 4. *Let $L\subseteq\Pi(E)$ be a simple lattice of permuting equivalences. Then any two prime quotients have the same set of indices. It consists of powers of p if E is finite and L contains a projective plane of order p as a sublattice.*

Jónsson [8] has represented the gluing of two Arguesian projective planes of different characteristics over a 2-element interval as a lattice of permuting equivalences on an infinite set — a finite set being impossible by the above. Moreover, we have

Corollary 5. *There is a finite lattice having a permuting equivalence representation which is not contained in the variety generated by all lattices of permuting equivalences on finite sets.*

Proof. Let p and q be different primes, L and $\hat{L}$ resp. the lattice which is the union of a projective 3-space $L_p=[0,b]$ of order p and $L_q=[a,1]$ of order q such that $ab\prec b$ and $a+b\succ a$ where $b\geqq a$ resp. $b\not\geqq a$. $\hat{L}$ is a subdirect product of L and a 2-element lattice, so it has a permuting equivalence representation this being the case for L according to JÓNSSON [8] — cf. the Appendix. On the other hand, in view of Corollary 4 we have one, hence both of $\varphi 0=\varphi b$ and $\varphi 1=\varphi a$ for every homomorphism φ of $\hat{L}$ into a lattice of permuting equivalences on a finite set. This gives rise to a separating identity since L is a projective modular lattice — as is well known.

Indeed, let φ be a homomorphism from a modular lattice M onto $\hat{L}$. Now, L_p and L_q are projective modular lattices according to FREESE [4], so we may choose preimages $L'_p \cong L_p$ and $L'_q \cong L_q$ in M and $[c')$ where c' is in L'_p with $\varphi c'=ab$. Let d' be in L'_q with $\varphi d'=a+b$, let b' be the top of L'_p and a' the bottom of L'_q. Then there is a sublattice L''_q of M mapped onto L_q, isomorphically, with bottom $a''=a'$ and $d''=b'd'+a'$ a point. Indeed, choosing a 4-frame of L'_q containing d' reduction with d'' yields a 4-frame of characteristic q generating L''_q — cf. [6], Corollary 3.4. Similarily, reduction with $b'd$ and $a'b'$ yields a 4-frame of characteristic p generating an isomorphic preimage L''_p of L_p such that $b'd/a'b'$ transposes down to b''/e'' where b'' is the top and e'' a plane of L''_p. Then, $L''_p \cup L''_q$ is a sublattice of M mapped onto L, isomorphically.

3. Von Neumann's Theorem

A ring R with 1 is regular if its principal left ideals form a complemented sublattice of the lattice of all left ideals or, equivalently, if every principal left ideal is generated by an idempotent — the same characterization is valid on the right. Equivalently, the finitely generated submodules of the left R-module ${}_RR^n$ form a complemented modular lattice $L_{fg}({}_RR^n)$ — see SKORNYAKOV [12], Chapter 2. According to von Neumann every complemented modular lattice with a spanning frame of order $n \geqq 4$ (or $n=3$ and Arguesian — JÓNSSON [9]) can be respresented in this way. But, such a lattice can be embedded into the subgroup lattice of an abelian group, firsthand, due to FRINK's Embedding Theorem [5], resp. JÓNSSON [8] — see also CRAWLEY—DILWORTH [1], Chapter 13. The frame can be chosen canonical, thus the following suffices for a proof of the Coordinatization Theorem. Of course, this approach uses the coordinatization of projective spaces.

Theorem 6. *Let A be an abelian group, L a complemented sublattice of $L(A^n)$, $n \geqq 3$, containing the canonical frame Φ with coordinate ring R. Then R is regular and*

$$\varphi(M) = \{(xr_1, \dots, xr_n) \mid x \in A,\ (r_1, \dots, r_n) \in M\}$$

defines an isomorphism of $L_{fg}({}_RR^n)$ onto L.

In this sense *every faithful representation of a complemented modular lattice with permuting spanning frame of order* $n \geqq 3$ *it obtained from the von Neumann—Jónsson coordinatization.*

Proof. The corollary follows from Theorem 1 and Claim 8 below. The proof of the theorem mimics, in the coordinatizing module, the calculations of VON NEUMANN [11] and DAY and PICKERING [2]. Write $\bar{r}$ for $(r_1, \ldots, r_n)$ and $\varphi\bar{r}$ for $\varphi(R\bar{r}) = \{(xr_1, \ldots, xr_n) \mid x \in A\}$. Let L' be the sublattice of L generated by Φ and R_{12}.

Claim 1. $\varphi(s\bar{r}) \leqq \varphi\bar{r}$, $\varphi(\bar{s}+\bar{r}) \leqq \varphi\bar{s}+\varphi\bar{r}$, so φ preserves joins.

Claim 2. $\varphi\bar{r} \in L$ if there is an i with $r_i = 0$, r_i invertible, or $r_i\bar{r} = \bar{r}$.

Proof. $\varphi\bar{r} = \bigcap_{j>1} \big((r_j)_{1j} + \sum_{k>1, k \neq j} a_k\big)$ for r_1 invertible,

$$\varphi\bar{r} = \big(\sum_{j>1} a_j\big) \cap \big(a_1 + \varphi(1, r_2, \ldots, r_n)\big) \quad \text{for} \quad r_1 = 0,$$

$$\varphi\bar{r} = \big(\varphi(r_1, 0, \ldots, 0) + \sum_{j>2} a_j\big) \cap \varphi(1, r_2, \ldots, r_n) \quad \text{for} \quad r_1\bar{r} = \bar{r}.$$

Claim 3. $R\bar{r} \subseteqq R\bar{s}$ iff $\varphi\bar{r} \subseteqq \varphi\bar{s}$ provided that $r_i = s_i = 0$ for an i.

Proof. Let $\varphi\bar{r} \subseteqq \varphi\bar{s}$ and $r_n = s_n = 0$. Then let

$$\bar{r}/\bar{s} = \{(x, y, 0, \ldots, 0) \mid x, y \in A, xr_i = -ys_i \text{ for } i < n\},$$

and note

$$\bar{r}/\bar{s} = \bigcap_{i<n} \big((r_i)_{1n} + (s_i)_{2n}\big) \cap (a_1 + a_2) \in L.$$

$\bar{r}/\bar{s} + a_2 \supseteqq a_1$ since for all x in A one has $-x\bar{r} \in \varphi\bar{r} \subseteqq \varphi\bar{s}$ which means $-x\bar{r} = y\bar{s}$ for a y in A. Now, let b be a complement of $a_2 \cap \bar{r}/\bar{s}$ in $[0, \bar{r}/\bar{s}]$. Then $b \in R_{12}$, so we have t in R with $b = t_{12} \subseteqq \bar{r}/\bar{s}$. But this implies that for every z in A there are x, y in A such that $z = x$, $-zt = y$, and $xr_i = -ys_i$ for $i < n$. Consequently, $zts_i = zr_i$ for all z and i. Since R consists of endomorphisms of A this means $ts_i = r_i$ and $t\bar{s} = \bar{r}$. So $R\bar{r} \subseteqq R\bar{s}$. The converse is clear.

Claim 4. For every $b \subseteqq a_i$ there is an idempotent r in R such that

$$b = \varphi(0, \ldots, \underset{i}{r}, \ldots, 0).$$

Proof. Let $i = 2$, d a complement of $(b + c_{12}) \cap a_1$ in $[0, a_1]$ and $e = d + (b + a_1) \cap c_{12}$. Then by modularity

$$e + a_2 \supseteqq d + (b + a_1) \cap (c_{12} + b) \supseteqq a_1,$$

$$e \cap a_2 = (b + a_1) \cap (d + c_{12}) \cap a_2 = (b + a_1 \cap a_2) \cap (d + c_{12}) = b \cap (d + c_{12}) =$$

$$= b \cap \big(d \cap (b + c_{12}) + c_{12}\big) = b \cap c_{12} = 0.$$

Therefore, $e\in R_{12}$ which means that $e=r_{12}$ for an r in R and

$$\varphi(0, r, 0, \ldots, 0) = (r_{12}+a_1)\cap a_2 = (a_1+b)\cap a_2 = b.$$

In addition, $a_1+c_{12}\cap e\supseteq a_1+(b+a_1)\cap c_{12}\supseteq b+a_1\supseteq e$, so $e=a_1\cap e+c_{12}\cap e$, and thus

$$r_{12} = a_1\cap r_{12}+c_{12}\cap r_{12} =$$

$$= \{(x, 0, \ldots, 0) \mid x\in A,\ xr = 0\}+\{(y, -y, 0, \ldots, 0) \mid y\in A,\ y = yr\}.$$

In other words, for every z in A there are x, y in A such that $z=x+y$, $zr=y$, $xr=0$, and $y=yr$, whence $zrr=zr$. Thus, $rr=r$.

Claim 5. R is a regular ring.

Proof. For r in R we have by Claim 4 an idempotent e with $\varphi(r, 0, \ldots, 0)=\varphi(e, 0, \ldots, 0)$ and $Rr=Re$ by Claim 3.

The following is shown in SKORNYAKOV [12], Chapter 2, § 5., Lemma 3.

Claim 6. For $M\subseteq R^k\times 0^{n-k}$ there is $\bar{s}$ with $s_k\bar{s}=\bar{s}$, $M=R\bar{s}+M\cap R^{k-1}\times\times 0^{n-k+1}$.

Claim 7. $M\subseteq R^m\times 0^{n-m}$ has generators $\bar{r}^{(1)}, \ldots, \bar{r}^{(m)}$ such that

$$r_k^{(k)}\bar{r}^{(k)} = \bar{r}^{(k)} \quad\text{and}\quad r_j^{(k)} = 0 \quad\text{for all}\quad k\leqq m,\quad k<j.$$

The proof is by induction on m using Claims 4 and 6.

Choosing such a generating set G for M we have $\varphi(M)=\sum_{\bar{r}\in G}\varphi\bar{r}$ by Claim 1. Thus by Claim 2 we have

Claim 8. $\varphi(M)$ belongs to L'.

Claim 9. Let $U\subseteq R^{k-1}\times 0^{n-k+1}$, $r_k\bar{r}=\bar{r}$, $s_k\bar{s}=\bar{s}$, and $r_j=s_j=0$ for all $j>k$. Then $\varphi\bar{r}\subseteq\varphi\bar{s}+\varphi(U)$ implies $\bar{r}\in R\bar{s}+U$.

Proof. Since $\varphi\bar{r}\subseteq\varphi\bar{s}+\varphi(U)$, we have for all x in A elements y, z in A and $\bar{t}$ in U such that $xr_kr_i=ys_ks_i+zt_i$ for all i. Since $t_j=0$ for $j\geqq k$ it follows $xr_k=xr_kr_k=ys_ks_k=ys_k$, $xr_kr_k=xr_ks_k$, and $xr_k(\bar{r}-\bar{s})=z\bar{t}$. Consequently, $\varphi(r_k(\bar{r}-\bar{s}))\subseteq\subseteq\varphi\bar{t}$ and $Rr_k(\bar{r}-\bar{s})\subseteq R\bar{t}$ by Claim 3, whence $R\bar{r}=Rr_k\bar{r}\subseteq Rr_k\bar{s}+R\bar{t}\subseteq R\bar{s}+U$.

Claim 10. $\varphi(N)\subseteq\varphi(M)$ implies $N\subseteq M$, so φ is one-to-one.

Let $M\subseteq R^k\times 0^{n-k}$ and proceed by induction on k. Let $\bar{r}^{(m)}\neq 0$ be a generator of N according to Claim 7. Then $\varphi\bar{r}^{(m)}\subseteq\varphi(N)\subseteq\varphi(M)$, $\bar{r}_j^{(m)}=0$ for $j>m$ and $m\leqq k$. If $m=k$ choose $\bar{s}$ by Claim 6 such that $M=R\bar{s}+U$ with $U=M\cap R^{k-1}\times\times 0^{n-k+1}$. Then $\bar{r}^{(m)}\in M$ by Claim 9. If $m<k$ let $U=(M+0^k\times R^{n-k})\cap(R^{k-1}\times 0^{n-k+1})$. Then $\varphi\bar{r}^{(m)}\subseteq\varphi(U)$, $\bar{r}^{(m)}\in U$ by the inductive hypothesis, and $\bar{r}^{(m)}$ is in M since $\bar{r}_j^{(m)}=0$ for $j\geqq k$.

Claim 11. φ is an onto map.

Proof. We show by induction on k that $a \leqq \sum_{i \leqq k} a_i$ is in the image. Let $f = \sum_{i<k} a_i$, $a \in L$, $a \leqq f + a_k$. Choose a complement b of $a+f$ in $[a, f+a_k]$ and c of $a \cap f$ in $[0, b]$. It follows

$$a \cap f + a \cap c = a \cap (a \cap f + c) = a \cap b = a,$$

$$c \cap f = c \cap b \cap (a+f) \cap f = c \cap a \cap f = 0,$$

$$c + f = c + a \cap f + f = b + f = b + a + f = f + a_k.$$

Now, $f = A^{k-1} \times 0^{n-k+1}$, $a_k = 0^{k-1} \times A \times 0^{n-k+2}$, and c is a subgroup of $f + a_k$. Thus, c defines a homomorphism of A into A^{k-1}, i.e.

$$c = \{(xr_1, \ldots, xr_{k-1}, x, 0, \ldots, 0)\}$$

for suitable r_i in R. Then there is a subgroup D of A such that

$$d = a \cap c = \{(xr_1, \ldots, xr_{k-1}, x, 0, \ldots, 0) \mid x \in D\},$$

$$(d+f) \cap a_k = \{(0, \ldots, 0, \underset{k}{x}, 0, \ldots, 0) \mid x \in D\} = \varphi(0, \ldots, \underset{k}{s}, \ldots, 0)$$

for an idempotent s in view of Claim 4. Let

$$\bar{t} = (sr_1, \ldots, sr_{k-1}, s, 0, \ldots, 0).$$

Then $s\bar{t} = \bar{t}$, $\varphi\bar{t} = \varphi(R\bar{t}) = d$. By the inductive hypothesis we have M' with $\varphi(M') = a \cap f$. Hence $\varphi(M' + R\bar{t}) = a$.

Addendum: Gluing of two representations

The proof of Lemma 3.5 in Jónsson [8] contains the following construction: Let L be the union of the ideal L_0 generated by b and the filter L_1 generated by a, $a \leqq b$. Let $\varphi_i \colon L_i \to \Pi(E_i)$ be representations for $i = 0, 1$, let $\alpha = \varphi_0 a$, $\beta = \varphi_1 b$. For each β-class X let a map γ_X of E_0 onto X with kernel α be given such that for all c in $[a, b]$ and x, y in E_0

$$x \;\varphi_0 c\; y \quad \text{if and only if} \quad \gamma_X x \;\varphi_1 c\; \gamma_X y.$$

Let $E = E_0 \times B$ where B is the set of β-classes. Define φc on E by

$$(x, X) \;\varphi c\; (y, Y) \quad \text{iff} \quad \begin{cases} X = Y \text{ and } x \;\varphi_0 c\; y & \text{for } c \leqq b \\ \gamma_X x \;\varphi_1 c\; \gamma_Y y & \text{for } c \geqq a. \end{cases}$$

Lemma 7 (Jónsson [8]). *φ is a representation of L on E which is permuting if and only if both φ_0 and φ_1 are permuting.*

Corollary 8. *There is a lattice not contained in any modular congruence variety but having a permuting equivalence representation on a finite set.*

Proof. For any prime p let $L_0=L(C^3_{p^2})$ and $L_1=L({}_RR^3)$ where $R=F_p[x]/x^2$, F_p the field of order p. Let $a=0\times C^2_{p^2}$, $b=R\times 0^2$, and L the lattice obtained by gluing L and L_1 over the 3-element chains $[a, 1_{L_0}]$ and $[0_{L_1}, b]$ to get $a=0_{L_1}\prec c\prec b=1_{L_0}$. Let φ_i: $L_i\to\Pi(E_i)$ be the canonical representations. To define the γ_X observe that E_0/α as well as each β-class has p^2 elements and is partitioned by $\varphi_i c$ into p classes of p elements. By the lemma we have a permuting equivalence representation on a p^{10}-element set. On the other hand L is not contained in any modular lattice variety generated by the congruence lattices of a class of algebraic structures closed under particular subdirect products, namely the congruences; see [13].

References

[1] P. Crawley and R. Dilworth, *Algebraic theory of lattices*, Prentice-Hall (Englewood Cliffs, New Jersey, 1973).

[2] A. Day and D. Pickering, The coordinatization of Arguesian lattices, *Trans. Amer. Math. Soc.*, **278** (1983), 507—522.

[3] J. Dénes and A. D. Keedwell, *Latin squares and their applications*, Academic Press (New York and London, 1974).

[4] R. Freese, Projective geometries as projective modular lattices, *Trans. Amer. Math. Soc.*, **251** (1979), 329—342.

[5] O. Frink, Complemented modular lattices and projective spaces of infinite dimensions, *Trans. Amer. Math. Soc.*, **60** (1946), 452—467.

[6] C. Herrmann, On the arithmetic of projective coordinate systems, *Trans. Amer. Math. Soc.*, **284** (1984), 759—785.

[7] C. Herrmann and A. Huhn, Lattices of normal subgroups which are generated by frames, in: *Lattice Theory* (Proc. Conf. Szeged, 1974), Colloq. Math. Soc. J. Bolyai, vol. 14, North-Holland (Amsterdam, 1976); pp. 97—136.

[8] B. Jónsson, Modular lattices and Desargues theorem, *Math. Scand.*, **2** (1954), 295—314.

[9] B. Jónsson, Representations of complemented modular lattices, *Trans. Amer. Math. Soc.*, **97** (1960), 64—94.

[10] B. Jónsson and G. Monk, Representation of primary Arguesian lattices, *Pacific J. Math.*, **30** (1969), 95—139.

[11] J. von Neumann, *Continuous geometry*, Princeton Univ. Press (Princeton, New Jersey, 1960).

[12] L. Skornyakov, *Complemented modular lattices and regular rings*, Oliver and Boyd (Edinburgh and London, 1964).

[13] R. Freese, C. Herrmann and A. P. Huhn, On some identities valid in modular congruence varieties, *Algebra Universalis*, **12** (1981), 322—334.

TECHNISCHE HOCHSCHULE DARMSTADT
FACHBEREICH MATHEMATIK
61 DARMSTADT, BRD

Acta Sci. Math., 51 (1987), 103—121

The diagram invariant problem for planar lattices

ROLAND JÉGOU, RICHARD NOWAKOWSKI and IVAN RIVAL

Dedicated to the memory of András P. Huhn

There are several graphical schemes in common use to represent a given finite ordered set. Of these the two that are best known are the 'comparability graph' and the 'diagram'.

The *comparability graph* of an ordered set P is the graph whose vertices are the elements of P and in which a pair of vertices x, y is adjacent if either $x<y$ or $x>y$. Much is known about the comparability graph: the characterization of comparability graphs (A. GHOUILA-HOURI [8], P. C. GILMORE and A. J. HOFFMAN [9]); the description of the order theoretical properties that are invariant among all ordered sets with the same comparability graph (M. HABIB [10]); the number of distinct orientations of a given comparability graph (L. N. ŠEVRIN and N. D. FILIPPOV [24]); a structure theory for comparability graphs (T. GALLAI [7]). (These and many further topics are treated closely in the recent survey articles of D. KELLY [13] and R. H. MÖHRING [17].)

The *diagram* of a finite ordered set P is that pictorial representation of P in the plane in which small circles, corresponding to the elements of P, are arranged in such a way that, for a and b in P, the circle corresponding to a is higher than the circle corresponding to b whenever $a>b$ and a straight line segment is drawn to connect the two circles whenever a covers b. Say that a *covers* b and write $a \succ b$ if $a>b$ and if $a>c\geqq b$ in P implies $c=b$. The diagram of P determines P up to isomorphism. Its economy of presentation accounts for the evident popularity of the diagram in the order literature today. Nevertheless, much less is known about it than about the comparability graph. (See I. RIVAL [23] for a recent survey of this theme.)

Closely related to the 'diagram' is the 'covering graph'. The *covering graph* of a finite ordered set P is the graph whose vertices are the elements of P and in which

Received April 24, 1986, and in revised form November 10, 1986.

a pair x, y of vertices is adjacent if $x \succ y$ or $y \succ x$. Not every graph is a covering graph and even one that is may have numerous 'orientations', that is, there may be many ordered sets with the same underlying covering graph. This article is inspired by the question, still little explored, of the order theoretical properties, if any, common to all of the ordered sets with the same covering graph. There seem to be few such properties. As a matter of fact, besides the trivial properties, such as the number of vertices or the number of edges we do not know even of one single property which remains invariant among all of the orientations of a fixed, but arbitrary, covering graph. Unlike the comparability graph none of these often studied numerical properties of an ordered set are such invariants: *width, length, dimension, jump number*. Indeed, there is even the intriguing possibility that there is *no* nontrivial diagram invariant at all!

Consider for example these common integer-valued functions defined on a finite ordered set P: the *width*

$$w(P) = \max \{|A| \mid A \text{ antichain in } P\};$$

the *length*

$$l(p) = \max \{|C|-1 \mid C \text{ chain in } P\};$$

the *dimension*

$$\dim P = \min \left\{m \mid L_1, L_2, \ldots, L_m \text{ linear extensions of } P \text{ and } \bigcap_{i=1}^{m} L_i = P\right\};$$

the *jump number*

$$s(P) = \min \{s(P, L) \mid L \text{ linear extension of } P\},$$

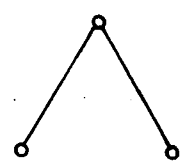
where

$$s(P, L) = \left|\{(a, b) \in P \times P \mid a \succ b \text{ in } L \text{ and } a \not\succeq b \text{ in } P\}\right|.$$

Suppose that P and P' are ordered sets with graph isomorphic comparability graphs. Then $w(P)=w(P')$, $l(P)=l(P')$, $\dim P = \dim P'$ and $s(P)=s(P')$ (cf. D. Kelly and W. T. Trotter [14], M. Habib [10]). Quite different is the situation for the covering graph — even for simple ordered sets. For instance, if $P \cong \mathbf{3}$ (see Figure 1) the three-element chain and $P' \cong 2\oplus 1$ (see Figure 2) then the covering graphs of P and P' are, of course, graph isomorphic. However, $w(\mathbf{3})=1<2=w(2\oplus 1)$, $l(\mathbf{3})=2>1=l(2\oplus 1)$, $\dim \mathbf{3}=1<2=\dim(2\oplus 1)$, and $s(\mathbf{3})=0<1=s(2\oplus 1)$. Actually the

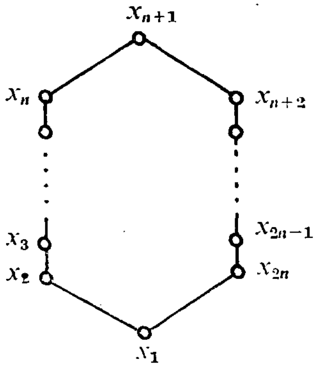
3

Figure 1



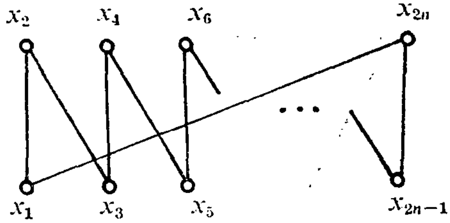
$2\oplus 1$

Figure 2



P_n

Figure 3



P_n'

Figure 4

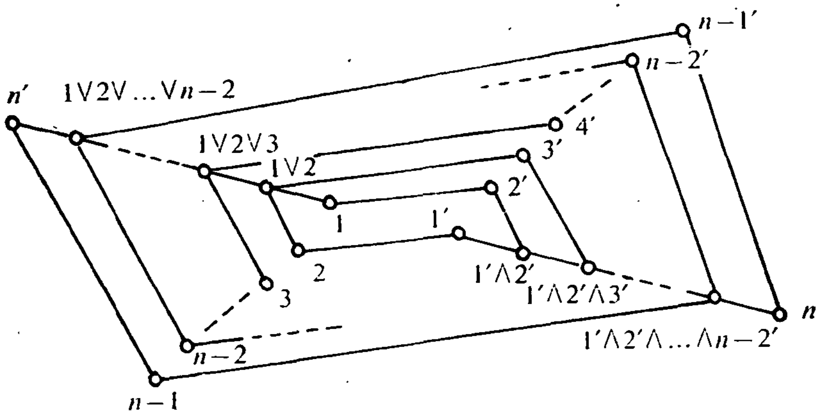
'deviation' can be much larger. For a positive integer n let P_n stand for the ordered set illustrated in Figure 3 and let P_n' stand for the $2n$-*cycle* illustrated in Figure 4. They have graph isomorphic covering graphs, yet $w(P_n)=2$, $w(P_n')=n$, $l(P_n)=n$, $l(P_n')=1$. $s(P_n)=1$, and $s(P_n')=n$. The dimension too differs, although for this pair of ordered sets the 'deviation' is only 1: $\dim P_n=2$ and $\dim P_n'=3$, provided that $n\geqq 3$. A more sophisticated example does show that the dimension can also 'deviate' considerably. A suitable pair of examples can be fashioned from an example constructed by D. Kelly [12] to show that planar ordered sets can have arbitrarily large dimension. Let $\mathbf{2}^n$ stand for the ordered set of all subsets of the n-element set. The subset $S_n=\{1, 2, ..., n\}\cup\{1', 2', ..., n'\}$ of the 'points' and 'copoints' $i'=\{1, 2, ..., i-1, i+1, ..., n\}$ of $\mathbf{2}^n$ has dimension n and so, in particular, also the subset of $\mathbf{2}^n$ consisting of

$$Q_n = S_n\cup\{1\vee 2\vee ...\vee i \mid 1\leqq i\leqq n-1\}\cup\{1'\wedge 2'\wedge ...\wedge i' \mid 1\leqq i\leqq n-1\}.$$

This ordered set Q_n of dimension n is illustrated in Figure 5 following the clever drawing of it proposed by D. Kelly [12]. A forty-five degree clockwise rotation of this illustration produces a planar lattice Q_n' of dimension two.

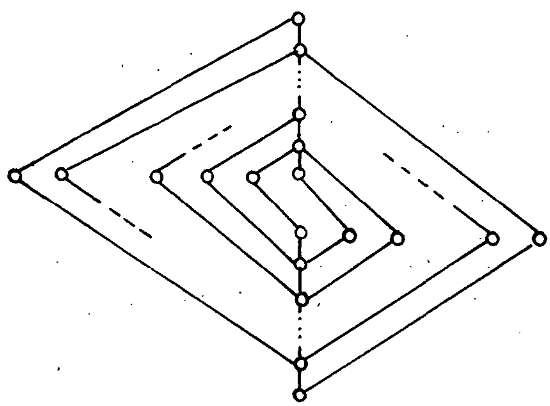
There is at least one residual positive fact. *Let P and P' be finite ordered sets with graph isomorphic covering graphs. If P is a chain then* $\dim P'-\dim P\leqq 1$. We must prove that $\dim P'\leqq 2$. To this end let $P=\{x_1\prec x_2\prec ...\prec x_n\}$, $n\geqq 3$, and suppose that P' is not a chain. There is no loss in generality if we assume too that x_1 is minimal in P'. We shall construct a chain decomposition $C_1, C_2, ..., C_m$, $m\geqq 2$, in this way. Let

$$C_1 = \{x_1 \prec x_2 \prec ... \prec x_i\},$$



Q_n

Figure 5



Q_n'

Figure 6

$1 \leqq i < n$, where x_{i+1} is maximal in P'. Thus, $x_{i+1} > x_i$. Suppose that $C_1, C_2, \ldots, C_t$ are already constructed and that $P - C_1 \cup C_2 \cup \ldots \cup C_t \neq \emptyset$. If

$$C_t = \{x_j \prec x_{j+1} \prec \ldots \prec x_k\}$$

where $j \leqq k$, then choose the least index $k+1 \leqq l \leqq n$ such that x_l is minimal in P' and set

$$C_{t+1} = \{x_l \prec \ldots \prec x_{k+2} \prec x_{k+1}\}.$$

If, on the other hand,

$$C_t = \{x_k \prec \ldots \prec x_{j+1} \prec x_j\},$$

with $j \leqq k$, then choose the least index $k+1 \leqq l \leqq n$ such that x_{l+1} is maximal in P' and, in this case, put

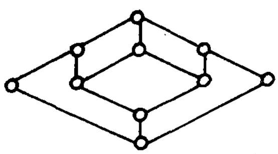
$$C_{t+1} = \{x_{k+1} \prec \ldots \prec x_{l-1} \prec x_l\}.$$

Finally, we can construct two linear extensions in two ways. First put

$$L_1 = C_1 \oplus C_2 \oplus \ldots \oplus C_m.$$

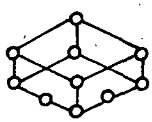
Now, construct L_2 in a similar way starting through with a 'dual' labelling beginning with x_n instead. It is easy to verify that $P' = L_1 \cap L_2$.

Our aim in this article is to consider a special case of this theme. Which order-theoretical properties are diagram invariants among all lattice orientations of a fixed, but arbitrary, covering graph of a planar lattice? Let P and P' be finite lattices with graph isomorphic covering graphs. If P is planar then must P' be planar too? The ordered sets illustrated in Figure 7 and Figure 8 show that this need not be the case at all. (Notice, moreover, that P' need not even be 'dismantlable' (cf. D. Kelly and I. Rival [14]).) It is a well known and useful fact that a planar lattice has, on either of its boundaries, elements which are both supremum irreducible and infimum irreducible, that is, *doubly irreducible* (cf. K. A. Baker, P. C. Fishburn and F. S. Roberts [2]). At least a fragment of this property is preserved by any lattice orientation.



P

Figure 7



P'

Figure 8

Theorem 1. *Let P and P' be finite lattices with graph isomorphic covering graphs. If P is planar then P' contains a doubly irreducible element too.*

Thus, the *existence* of a doubly irreducible element is a diagram invariant among all lattice orientations of a planar lattice. To prove this result we shall make extensive use of the geometric theory of planarity and planar embeddings for finite lattices established in D. Kelly and I. Rival [14] and in C. R. Platt [20].

Theorem 2. *Let P and P' be finite planar lattices with graph isomorphic covering graphs. If for some planar embedding $e(P)$ of P every doubly irreducible element of P lies on the boundary of $e(P)$ then, for any planar embedding $e(P')$ of P', the set of faces of $e(P')$ equals the set of faces of $e(P)$.*

Thus, under these hypotheses, the *set* of faces of any planar embedding of any planar lattice orientation of a planar lattice is a diagram invariant.

The problem to determine the lattice re-orientations of a fixed, but arbitrary, covering graph of a lattice has been more extensively studied, especially for distributive and modular lattices (cf. L. ALVAREZ [1], J. JAKUBÍK [11], D. DUFFUS and I. RIVAL [5], I. RIVAL [23]).

Planarity

The purpose of this section is to clarify the basic terms which we require to prove Theorem 1 and Theorem 2. The important references for our point of view are O. ORE [19], D. KELLY and I. RIVAL [15] and C. R. PLATT [20].

A finite graph G is *planar* if it can be embedded in the plane R^2 using a Jordan arc (that is, a homeomorphic image of the closed unit interval) for each edge such that different edges have, at most their endpoints in common. We denote by $e(G)$ such a *planar realization* of G. A *simple Jordan curve* or, for brevity, a *Jordan curve* in R^2 is a homeomorphic image of the unit circle. According to the well known Jordan Curve Theorem any Jordan curve C partitions the rest of the plane into two open sets, the *interior* $\mathrm{Int}\, C$ of C, and the *exterior* $\mathrm{Ext}\, C$ of C. Any Jordan arc connecting two vertices in $e(G)$ corresponds to an (elementary) *path* of G. Similarly, any Jordan curve in $e(G)$ corresponds to an (elementary) cycle of G.

We shall apply the Jordan Curve Theorem in this form. *Let C be a Jordan curve, let $x \in \mathrm{Ext}\, C$, and let $y \in \mathrm{Int}\, C$. Then any Jordan arc connecting x and y meets C in at least one point.*

To each planar realization $e(G)$ of G we associate a set of (connected) domains $\{F_0, F_1, \ldots, F_k\}$ in R^2 called the *faces* of $e(G)$. (For a definition of a 'face' see, for example, O. ORE [19].) There is just one unbounded domain F_0, the exterior face of $e(G)$; the other domains defining the interior faces satisfy $\mathrm{Int}\, F_i \cap e(G) = \emptyset$, $1 \leqq i \leqq k$, where $\mathrm{Int}\, F_i$ stands for the topological interior of the domain F_i.

Let w and y be two distinct points of a Jordan curve C. There are exactly two Jordan arcs lying in C having only w and y in common, say $A(w, y)$ and $A(y, w)$. Four distinct points of C constitute a *quadrilateral* (w, x, y, z) on C if $x \in A(w, y)$ and $z \in A(y, w)$. This basic topological property is due to C. R. PLATT [20]. *Let C be a Jordan curve and let (w, x, y, z) be a quadrilateral on C. Let E (respectively F) be a Jordan arc with endpoints w and y (respectively x and z) and suppose that E and F are both outside or both inside C. Then $E \cap F \neq \emptyset$.*

We treat now some of the basic terminology concerning 'planar' ordered sets as developed in D. KELLY and I. RIVAL [15]. Let P be a finite ordered set. Let π_1 and π_2 stand for the first and the second projections of R^2 onto R. A *planar embedding* $e(P)$ of P consists of

(1) an injection $x \to \bar{x}$ from P to R^2 such that $\pi_2(\bar{x}) < \pi_2(\bar{y})$ whenever $x < y$ in P, and

(2) straight line segments $\bar{x}\bar{y}$ connecting $\bar{x}$ and $\bar{y}$ whenever $x \prec y$ in P, and which do not intersect except possibly at their endpoints.

For simplicity of notation we shall identify each point $\bar{x}$ in the plane with the corresponding point x in P and use x for both. P is *planar* if it has a planar embedding. A planar representation $e(P)$ of P is defined by (1) above and

(2)′ increasing Jordan arcs denoted by xy with endpoints x and y whenever $x \prec y$ in P, and which do not intersect except possibly at their endpoints.

An increasing Jordan arc is defined by $A = \{(f(t), t) \mid t \in [\alpha, \beta]\}$, where f is a continuous function from a closed interval $[\alpha, \beta]$ of R to R. A *decreasing* Jordan arc is defined similarly.

D. Kelly has proved that these two notions are equivalent, so we may speak by turns of planar embeddings and planar representations. A maximal chain from x to y in P, denoted by (x, y), is a sequence $x = x_0, x_1, \ldots, x_n = y$ of elements of P with $x_i \prec x_{i+1}$, $0 \leqq i \leqq n-1$. In a planar representation $e(P)$ of a planar ordered set P any increasing Jordan arc in $e(P)$ connecting two vertices x and y and denoted by $A^+(x, y)$, corresponds to a maximal chain (x, y) of P.

Let G be a finite graph. We shall denote by $\mathscr{P}(G)$, $\mathscr{L}(G)$ and $\mathscr{L}_P(G)$ the sets of all ordered sets, all lattices and all planar lattices, respectively, each of whose covering graphs is G. Obviously $\mathscr{L}_P(G) \subseteq \mathscr{L}(G) \subseteq \mathscr{P}(G)$. G is called *orientable* whenever $\mathscr{P}(G) \neq \emptyset$. If G is any connected orientable graph G having at least two edges then $\mathscr{L}(G) \subseteq \mathscr{P}(G)$. On the other hand, if $\mathscr{L}_P(G) \neq \emptyset$ the equality $\mathscr{L}(G) = \mathscr{L}_P(G)$ need not hold at all (cf. Figure 7 and Figure 8).

Let G be a graph and suppose that $\mathscr{L}_P(G) \neq \emptyset$. Let $L \in \mathscr{L}_P(G)$ and let $e(L)$ be a particular planar realization of G. Let

$$F(e(L)) = \{F_0, F_1, \ldots, F_k\}$$

denote the set of faces of $e(L)$. It is a trivial consequence of the familiar Euler formula relating the numbers of vertices, edges and faces that *the number of faces is an invariant of* $\mathscr{L}_P(G)$, that is, if $L, L' \in \mathscr{L}_P(G)$ and $e(L)$, $e(L')$ are corresponding planar representations then

$$|F(e(L))| = |F(e(L'))|.$$

The vertices corresponding to the (not necessarily elementary) cycle of G associated with the exterior face F_0 of a planar representation $e(L)$ of $L \in L_P(G)$ determine the *boundary* $B(e(L))$ of $e(L)$. We can define the *left boundary* and the *right boundary* as the maximal chains corresponding to the Jordan arcs $A_l^+(0, 1)$ and

$A_r^+(0, 1)$ of $B(e(L))$ connecting the images of the extremal elements of L such that for each $x \in A_l^+(0, 1)$ and for each $y \in A_r^+(0, 1)$ satisfying $\pi_2(x) = \pi_2(y)$ then

$$\pi_1(x) \leqq \pi_1(y).$$

A *region* of $e(L)$ is a subset of L consisting of all elements of L in the area of the plane bounded by the Jordan arcs corresponding to the maximal chains C and D having the same extremal elements. A subset S of an ordered set P is *cover-preserving* if $x \prec y$ in S implies $x \prec y$ in P. *A region is a planar cover-preserving sublattice of* L (D. KELLY and I. RIVAL [15]).

Call a region a *strict region* if it is defined by two maximal chains having only their extremal elements in common. Therefore, the geometric curve in $e(L)$ corresponding to such a region R is a Jordan curve consisting of two increasing Jordan arcs having only their extremal elements in common. These endpoints are the images in $e(L)$ of the least and greatest elements of R. *Any interior face of* $e(L)$ *is a strict region of* L *whose interior contains no vertices or edges of* L (D. KELLY and I. RIVAL [15]).

In what follows we assume that, for $L \in \mathscr{L}_P(G)$ and for one of its embeddings $e(L)$, $B(e(L))$ is a strict region. Otherwise, L can be decomposed into a linear sum $L_1 \oplus L_2 \oplus \ldots \oplus L_k$ in which the top element of L_i is the bottom element of L_{i+1}, for each $i = 1, 2, \ldots, k-1$. In that case we can apply this more general result.

Proposition 3. *If* $L \in \mathscr{L}(G)$ *is a linear sum*

$$L = L_1 \oplus \ldots \oplus L_k$$

then, for any $L' \in \mathscr{L}(G)$,

$$L' = L_1' \oplus \ldots \oplus L_k' \quad or \quad L'^d = L_1' \oplus \ldots \oplus L_k',$$

where each L_i *and* L_i' *have the same covering graph and the same extremal elements.*

Proof. It is enough to prove this property with $L = L_1 \oplus L_2$. Let us denote by a the greatest element of L_1; it is also the least element of L_2. Let $L' \in \mathscr{L}(G)$. The element a cannot be either the least element $0'$ of L' or the greatest element $1'$ of L'. For, if $a = 0'$, say, then we may consider the element $0 \vee 1$ in L', where 0 and 1 are the least and greatest elements of L. Using maximal chains $(0, 0 \vee 1)$ and $(1, 0 \vee 1)$ in L' we can construct a path from 0 to 1 in L' which does not contain a. This is a contradiction, since every maximal chain in L from 0 to 1 must contain a.

We can suppose that there exists x in L satisfying $x < a$, in L and $x < a$ in L', for otherwise we consider L'^d, the dual of L'.

Let y be any element of $L_2 - \{a\}$, that is, $y > a$. There is at least one maximal chain in L from x to y and all such maximal chains contain a. Consequently, any path from x to y in L' contains a. We claim that $y > a$ in L'. Otherwise, consider $x \wedge y$ in L'. There exist two maximal chains in L' not containing a, $(x \wedge y, x)$ and

$(x \wedge y, y)$. Then we can construct in L' a path from x to y not containing a either. That is a contradiction.

Similarly, if $z < a$ in L then $z < a$ in L'. This completes the proof.

Now, we consider $L \in \mathscr{L}_P(G)$, any one of its planar embeddings $e(L)$ and $F \in F(e(L))$. For any $L' \in \mathscr{L}(G)$, let $D'(F)$ denote the subdiagram of L' induced by the elementary cycle of G corresponding to F.

Proposition 4. *Let $L \in \mathscr{L}_P(G)$, let $e(L)$ be a planar embedding of L, let $F \in F(e(L))$ and let $L' \in \mathscr{L}(G)$. Then $D'(F)$ is a planar lattice.*

Proof. We must show that $D'(F)$ consists of two maximal chains of L'. For contradictions suppose not. Then $D'(F)$ has at least two maximal and at least two minimal elements. ($D'(F)$ is an elementary cycle.) Let w and y be two distinct minimal elements of $D'(F)$ such that $h(w) \leqq h(y)$ where h is the height function of L' and let x, z be the two elements on the cycle adjacent to y.

There exist four maximal chains $(0', w)$, $(0', y)$, $(x, 1')$ and $(z, 1')$ in L', where $0'$, $1'$ are, respectively, the bottom, top elements of L' satisfying

$$(0', w) \cap (x, 1') = \emptyset, \quad (0', w) \cap (z, 1') = \emptyset,$$

$$(0', y) \cap (x, 1') = \emptyset \quad \text{and} \quad (0', y) \cap (z, 1') = \emptyset.$$

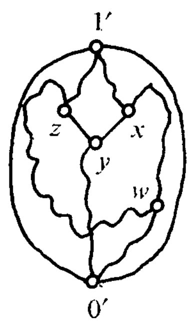


Figure 9

Using the Jordan arcs corresponding to these chains in $e(L)$ we can easily construct two Jordan arcs $A(w, y)$ and $A(x, z)$ of $e(L)$ which are inside C (the Jordan curve corresponding to F in $e(L)$) if F is the exterior face, or outside, otherwise (Fig. 10).

Now, (w, x, y, z) forms a quadrilateral on C and that is a contradiction since then $A(w, y)$ and $A(x, z)$ intersect in a point which is not a vertex of L. Then $D'(F)$ is formed by two maximal chains of L having in common only their endpoints.

In the case that $L' \in \mathscr{L}_P(G)$ we derive this consequence.

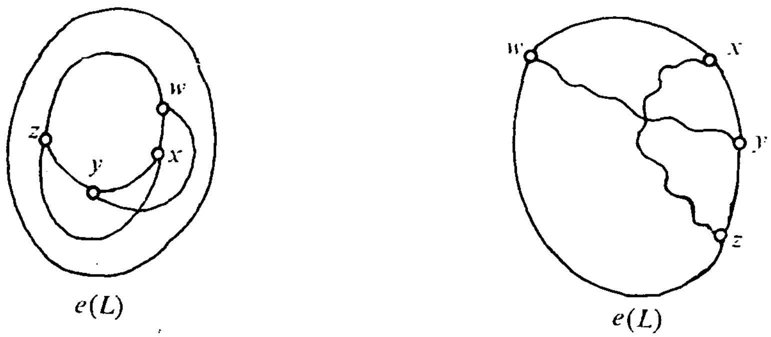


Figure 10

Corollary 5. *Let* $L, L' \in \mathscr{L}_P(G)$ *and let* $e(L)$, $e(L')$ *be corresponding planar representations. Then any face of* $e(L)$ *is a strict region of* $e(L')$.

The subset corresponding to a face of $e(L)$ is a planar, cover-preserving sublattice of L and is transformed to a planar, cover-preserving sublattice of L'. But this does not hold if we consider any planar, cover-preserving sublattice of L. The next figure, Figure 11, illustrates an example in which a sublattice of L is not transformed to a sublattice of L'.

Figure 11

Irreducible elements

An element of a lattice L is *doubly irreducible* in L if it has at most one lower cover and at most one upper cover. Let Irr (L) denote the set of all doubly irreducible elements of L. This fact was in a sense the start of the theory of planar lattices. *Any planar lattice has at least one doubly irreducible on the left boundary of any of its planar embeddings* (K. A. Baker, P. C. Fishburn, and F. S. Roberts [2]).

Proposition 6. *If* L *and* L' *are lattice orientations of the same covering graph* G, *then*

$$\big||\mathrm{Irr}(L)| - |\mathrm{Irr}(L')|\big| \leqq 2.$$

Proof. Let us consider $T(G)=\{x\in V \mid \deg(x)=2\}$ where V is the vertex set of G and $\deg x$ is the *degree* of x. Then $\mathrm{Irr}(L)\subseteqq T(G)$ and

$$|T(G)|-2 \leqq |\mathrm{Irr}(L)| \leqq |T(G)|$$

because the greatest and the least elements of L can be in $T(G)$.

The number of doubly irreducible elements is not an invariant. In Figure 12 we illustrate three planar lattice orientations of the same covering graph having respectively 3, 4 and 5 doubly irreducible elements.

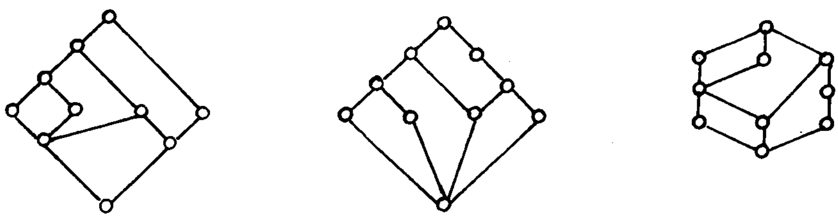


Figure 12

We are ready to prove our first principal result concerning doubly irreducible elements.

Theorem 1. *If* $\mathscr{L}_P(G)\neq\emptyset$ *then for each* $L'\in\mathscr{L}(G)$, $\mathrm{Irr}(L')\neq\emptyset$.

Proof. Consider $L\in\mathscr{L}_P(G)$ and let $e(L)$ be one of its planar embeddings. If $|\mathrm{Irr}(L)|\geqq 3$ the result is obvious according to the Proposition. Hence we can suppose that $\mathrm{Irr}(L)=\{a_1, a_2\}$ with a_1 on the left boundary of $e(L)$, say.

Now consider $L'\in\mathscr{L}(G)$ and suppose that $\mathrm{Irr}(L')=\emptyset$. Then L' cannot be planar and a_1 and a_2 must be the least and the greatest elements of L'. Our aim now will be to construct a planar embedding of L', which is a contradiction.

Let F_1 be the face containing a_1 in L and let us use F_1 too to denote the corresponding path in L'. According to Proposition 4, F_1 can have only one maximal element in L'.

Suppose that we have constructed in L' the subset corresponding to the faces $F_1, F_2, \dots, F_k$, $k\geqq 1$, of L. $F_1\cup\dots\cup F_k$ is a planar subset of L. Let us denote by $B_k(L)$ the path corresponding to its boundary, with respect to $e(L)$.

We shall show that the subset 'generated' by $\{F_1, \dots, F_k\}$ in L' is planar and its boundary $B_k(L')$ is exactly the path $B_k(L)$.

The assertion is true for $k=1$. We proceed by induction on k. Let us denote by S the set of vertices of $B_k(L)$ which are contained in new faces of $e(L)$ in the sense that each vertex of S is adjacent to a vertex in $\mathrm{Ext}\, B_k(L)$. Let x_1 be a vertex of S of minimum height in L'. If F_{k+1} is a new face containing x_1 in L we can write

$$F_{k+1} = \{x_1, x_2, \dots, x_n, x_{n+1}, \dots, x_m\}$$

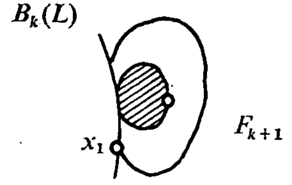


Figure 13

such that $n \geqq 3$, $m \geqq n$, $\{x_n, x_{n+1}, \ldots, x_m, x_1\} \subseteq B_k(L)$ and $x_2, \ldots, x_{n-1} \notin B_k(L)$ (and possibly $x_n = x_m = x_1$). Notice that, since $\operatorname{Irr}(L') = \emptyset$, there can be no face within this new face F_{k+1} (cf. Figure 13).

Then the subset of L' generated by F_{k+1} has exactly one maximal element x_i and one minimal element x_j such that $1 \leqq i \leqq n$ and, $n \leqq j \leqq m$ or $j = 1$.

Now $x_2 > x_1$ in L'. Otherwise there exists a maximal chain $(a_1, \ldots, x_2, x_1)$ from a_1 to x_1 in L'. If y is the last point of intersection of this chain with $B_k(L') = B_k(L)$ then $y \in S$ and we have a contradiction because $h(y) < h(x_2) < h(x_1)$ in L', where h is the height function of L'.

On the other hand, if $x_n \neq x_1$ we have $x_{n+1} < x_n$ in L'. Indeed, if $x_{n+1} > x_n$ then consider three maximal chains (a_1, x_1), (a_1, x_n) and (x_{n+1}, a_2). Since $x_n \in S$, $h(x_1) \leqq h(x_n) < h(x_{n+1})$, $(a_1, x_1) \cap (x_{n+1}, a_2) = \emptyset$ and also $(a_1, x_n) \cap (x_{n+1}, a_2) = \emptyset$.

In L, using the Jordan arcs $A(a_1, x_1)$ and $A(a_1, x_n)$ corresponding to (a_1, x_1) and (a_1, x_n), respectively, we can construct a Jordan arc $A(x_1, x_n)$.

If $a_2 \notin F_{k+1}$ then the Jordan curve $C = A(x_1, x_n) \cup A(x_1, x_2, \ldots, x_n)$ (the latter part corresponding to the path $(x_1, x_2, \ldots, x_n)$ of F_{k+1}), has x_{n+1} in its inside which is a contradiction since $A(x_{n+1}, a_2)$, the Jordan arc associated with (x_{n+1}, a_2) would cut C in a point which is not a vertex of L (see Figure 14).

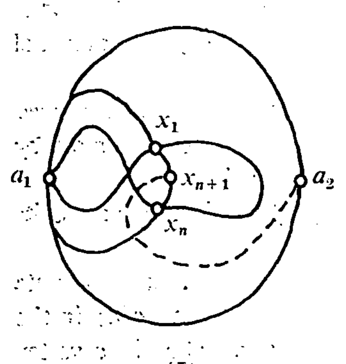


Figure 14

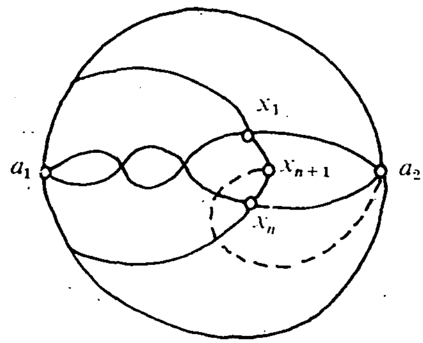


Figure 15

If $a_2 \in F_{k+1}$ then (x_1, a_2, x_n, x_{n+1}) is a quadrilateral on F_{k+1} since $x_1 \neq a_2$ and $x_n \neq a_2$. The previous Jordan arcs $A(x_1, x_n)$ and $A(x_{n+1}, a_2)$ must cut outside F_{k+1}, which is contradiction (cf. Figure 15).

We conclude by verifying that $B_{k+1}(L) = (B_k(L) - \{x_{n+1}, \ldots, x_m\}) \cup \cup \{x_2, \ldots, x_{n-1}\} = B_{k+1}(L')$.

So, if $F_1 \cup \ldots \cup F_k$ has a planar embedding in L', then $F_1 \cup \ldots \cup F_{k+1}$ also has a planar embedding in L'. Thus we obtain a planar embedding of L'.

The converse of Theorem 1 is false. There are finite graphs G such that, for each $L' \in \mathscr{L}(G)$, $\mathrm{Irr}(L') \neq \emptyset$, and yet there is no planar lattice orientation of G at all (see Figure 16).

Figure 16

Faces

Let G be a graph such that $\mathscr{L}_P(G) \neq \emptyset$ and let $L, L' \in \mathscr{L}_P(G)$. We denote by 0, 1 and $0'$, $1'$ the extremal elements of L and L', respectively. In this section we consider relations that exist between the faces of planar representations $e(L)$ and $e(L')$, of L and L'.

We shall require this.

Lemma 7. *If there exists L in $\mathscr{L}_P(G)$ such that $\mathrm{Irr}(L) \subseteq B(e(L))$ for a planar embedding $e(L)$ of L then $B(e(L)) = B(e(L'))$ for any planar embedding $e(L')$ of any L' in $\mathscr{L}_P(G)$.*

Proof. According to the Proposition 3 we can suppose that the exterior face F_0 of $e(L)$ is a strict region of L. Hence this property is also true for L'. Using the Corollary 5 the Jordan curve C_0' corresponding to the image of F_0 in L' defines a strict region, so it is a planar, cover-preserving sublattice of L'.

Let us assume that $B(e(L')) \neq B(e(L))$. By hypothesis, there exist at least two elements $a, b \in B(e(L)) \cap B(e(L'))$ such that $a, b \in \mathrm{Irr}(L')$, a is on the left boundary of L' and b is on the right boundary of L'. Let us denote by x, y, z, t respectively the four elements on the path of L' corresponding to $A(a, b) \subseteq C_0'$ such that $1 \in A(a, b)$, y is the first element not in $B(e(L'))$, x is its predecessor, t is the first element after y on $B(e(L'))$ and z is its predecessor, if the path is directed from a

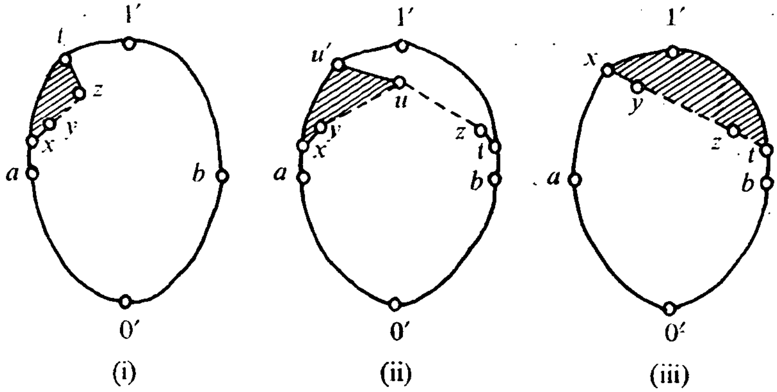


Figure 17

to b. In this way we get three different configurations as illustrated schematically in Figure 17. In the case (ii) u is the greatest element of the sublattice defined by C_0' in L' and u' is the first common element of a maximal chain from u to $1'$ in L'.

In each of these three cases the shaded strict regions must be planar, cover-preserving sublattices of L' and thus have at least one doubly irreducible element on their left boundaries, for cases (i) and (ii), and on the right boundary too. Hence we obtain a contradiction because the doubly irreducible elements of L' are necessarily on C_0'. Therefore $B(e(L'))=B(e(L))$.

Theorem 2. *If there exists* L *in* $\mathscr{L}_P(G)$ *such that* $\operatorname{Irr}(L)\subseteqq B(e(L))$ *for some planar embedding* $e(L)$ *of* L *then any planar embedding* $e(L')$ *of any* $L'\in\mathscr{L}_P(G)$ *satisfies* $F(e(L'))=F(e(L))$.

Proof. The previous lemma implies the invariance of the exterior face and thus $\operatorname{Irr}(L')\subseteqq B(e(L'))$.

Let F be any interior face of L. In L' the image of F defines a planar cover-preserving sublattice, say $L'(F)$. Let $0_F'$ and $1_F'$ denote its least and greatest elements, respectively.

Let us suppose that $L'(F)$ is not a face of $e(L')$.

If x and y are, respectively, elements on the left and on the right boundaries of $L'(F)$ then we claim that every path in $L'(F)$ from x to y contains either $0_F'$ or $1_F'$. Indeed if such a path $p(x,y)$ does not exist, let $(1_F', 1')$ and $(0', 0_F')$ be two maximal chains from $1_F'$ to $1'$ and from $0'$ to $0_F'$, respectively, in L'. We have $(0', 0_F')\cap\cap p(x,y)=\emptyset$ and $(1_F', 1')\cap p(x,y)=\emptyset$.

Now consider $A(x,y)$, $A(1_F', 1')$ and $A(0', 0_F')$, the Jordan arcs in $e(L)$ corresponding to these paths. We know that $0'$ and $1'$ lie on the boundary of L (because $B(e(L))=B(e(L'))$. Thus we can consider $A(0', 1')$, a Jordan arc connecting $0'$ and $1'$ and having only these two points in common with $e(L)$. Therefore

$$A(0'_F, 1'_F) = A(0'_F, 0') \cup A(0', 1') \cup A(1', 1'_F) \quad \text{and} \quad A(x, y)$$

are two Jordan arcs lying outside of the Jordan curve C corresponding to F in $e(L)$.

Now, using the quadrilateral $(0'_F, y, 1'_F, x)$ on C these two Jordan arcs must cross at a point which is not a vertex of L. That is a contradiction (see Figure 18).

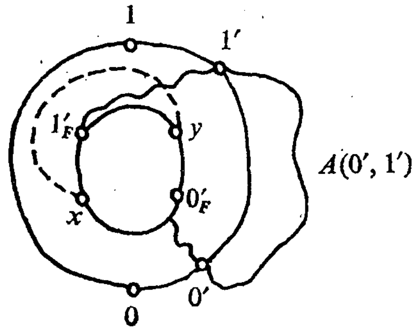


Figure 18

Now consider the upper covers $x_1, x_2, \ldots, x_k$, $k \geqq 2$, of $0'_F$ in $L'(F)$. In $L'(F)$, $\{x_i, x_{i+1}\}$, $1 \leqq i < k$, are in the same face, say F_i. Then there is $1 \leqq i \leqq k-1$ such that $1'_F \in F_i$. For otherwise, for each $1 \leqq i \leqq k-1$, $1'_F \notin F_i$ and there is a path from x_1 to x_k containing neither $0'_F$ nor $1'_F$, a contradiction according to the previous property.

The region R defined by the left boundaries of $L'(F)$ and F_i (see Figure 19) is a planar sublattice of L'. Then it must have at least one doubly irreducible element on its right boundary, which is impossible because $\mathrm{Irr}\,(L') \subseteq B(e(L'))$.

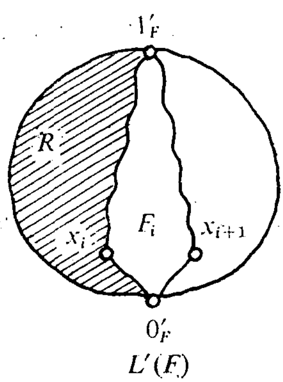


Figure 19

Then $L'(F)$ cannot contain an element in its inside and it must be a face of $e(L')$.

The converse of Theorem 2 is false. The planar lattice illustrated in Figure 20 has essentially just one planar lattice orientation and, in particular, the set of faces is invariant. Nevertheless, not all of its doubly irreducible elements lie on the boundary.

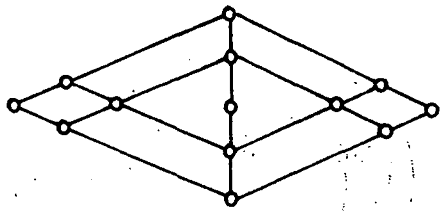
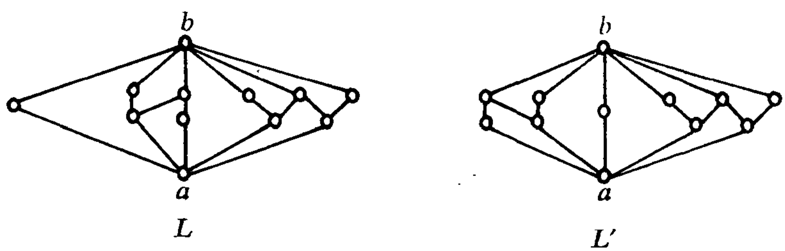


Figure 20

A conjecture

An understanding of the re-orientations of planar lattices may well advance our knowledge of the orientations of covering graphs. Are there 'canonical operations' which 'transform' one planar lattice orientation to another?

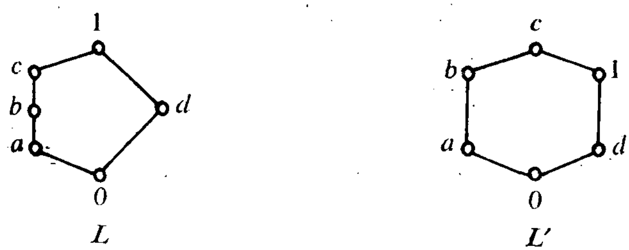
D. Kelly and I. Rival [15] have described a procedure, call it *permutation-reflection*, which can be applied to produce all planar embeddings from any fixed planar embedding of a planar lattice. Loosely speaking the idea is to consider $a, b \in L$ such that $a < b$ and all regions $R_1, R_2, \ldots$ with a and b as extremal elements. If $R_i \cap R_j = \{a, b\}$, we permute $R_i - \{a, b\}$ with $R_j - \{a, b\}$ (according to the linear order defined by considering the projections on the x-axis), without affecting the planarity itself (cf. Figure 21). Every planar embedding of L



L' is obtained from L by permutation-reflections
Figure 21

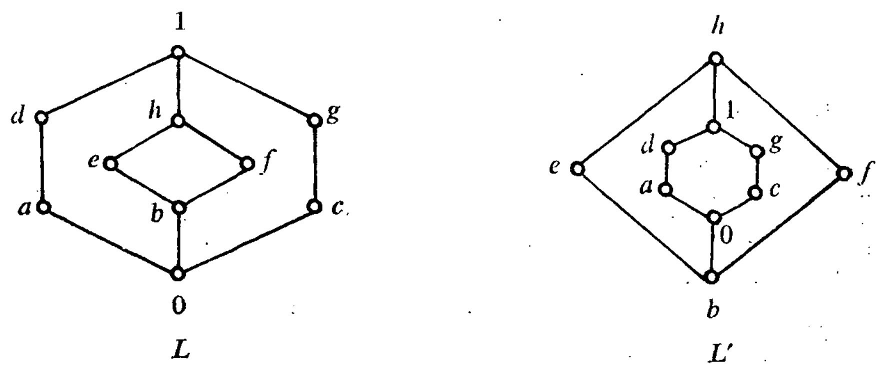
is produced from any fixed one by a sequence of permutation-reflection transformations.

For $L, L' \in \mathscr{L}_P(G)$ we say that L' is obtained from L by a *rotation* of L provided there are planar embeddings $e(L)$, $e(L')$ of L, L', respectively, such that $F(e(L))=F(e(L'))$. For instance, under the hypotheses of Theorem 2, every planar lattice orientation L' is obtained from L by a rotation (cf. Figure 22).



L' is obtained from L by a rotation
Figure 22

The stereographic projection from the sphere to the plane and its inverse obtained by selecting the north pole in some face F produces a different planar embedding of a planar graph with F as its exterior face (cf. O. ORE [19], C. R. PLATT [20]). This transformation which we shall call *inversion* leaves fixed the set of faces (see Figure 23). This transformation applied to an arbitrary face of a planar embedding of a planar lattice will not necessarily produce another planar lattice embedding. We do not at this time yet know which faces of a planar lattice embedding can be the exterior faces of a planar lattice embedding, obtained by inversion.



L' is obtained from L by inversion
Figure 23

Here is our conjecture. *Any planar embedding* $e(L')$ *of any* $L' \in \mathcal{L}_P(G)$ *is obtained from any* $L \in \mathcal{L}_P(G)$ *by a sequence of transformations each either a permutation-reflection, or a rotation, or an inversion.*

Added in proof. Theorem 1 has an important extension to *dismantlable lattices* (cf. D. Kelly and I. Rival [14], [15]).

Corollary. *Let* P *and* P' *be finite lattices with graph isomorphic covering graphs. If* P *is dismantlable then* P' *contains a doubly irreducible.*

Proof. If P is planar then the assertion is precisely Theorem 1. If P is non-planar then the dimension of P is at least three (cf. [15]). In this case P contains at least three doubly irreducible elements (cf. Theorem 6.11, [15]). Then, as in Proposition 6, P' must contain a doubly irreducible element as well.

References

[1] L. Alvarez, Undirected graphs as graphs of modular lattices, *Canad. J. Math.*, **17** (1965), 923—932.

[2] K. A. Baker, P. C. Fishburn and F. S. Roberts, Partial orders of dimension 2, *Networks*, **2** (1971), 11—28.

[3] M. Chein and P. Martin, Sur le nombre de sauts d'une forêt, *C. R. Acad. Sci. Paris*, **275** (1972), 159—161.

[4] M. Chein and M. Habib, The jump number of dags and posets: an introduction, *Ann. Discrete Math.*, **9** (1980), 189—194.

[5] D. Duffus and I. Rival, Path length in the covering graph of a lattice, *Discrete Math.*, **19** (1977), 139—158.

[6] B. Dushnik and E. W. Miller, Partially ordered sets, *Amer. J. Math.*, **63** (1941), 600—610.

[7] T. Gallai, Transitiv orientierbare Graphen, *Acta Math. Acad. Sci. Hungar.*, **18** (1967), 25—66.

[8] A. Ghouila-Houri, Caractérisation des graphes non orientés dont on peut orienter les arêtes de manière à obtenir le graphe d'une relation d'ordre, *C. R. Acad. Sci., Paris*, Sér. A—B, **254** (1962), 1370—1371.

[9] P. C. Gilmore and A. J. Hoffman, A characterization of comparability graphs and of interval graphs, *Canad. J. Math.*, **16** (1964), 539—548.

[10] M. Habib, Comparability invariants, in: *Orders: Descriptions et Rôles* (eds. M. Pouzet and D. Richard), North-Holland, (Amsterdam 1984); pp. 371—386.

[11] J. Jakubík, Unoriented graphs of modular lattices, *Czechoslovak Math. J.*, **25** (1975), 240—246.

[12] D. Kelly, Fundamentals of planar ordered sets, *Discrete Math.*, to appear.

[13] D. Kelly, Comparability graphs, in: *Graphs and Order* (I. Rival, ed.), D. Reidel (Dordrecht, 1985); pp. 3—40.

[14] D. Kelly and I. Rival, Crowns, fences, and dismantlable lattices, *Canad. J. Math.*, **26** (1974), 1257—1271.

[15] D. Kelly and I. Rival, Planar lattices, *Canad. J. Math.*, **27** (1975), 636—665.

[16] D. Kelly and W. T. Trotter, Dimension theory for ordered sets, in: *Ordered sets* (I. Rival, ed.), D. Reidel (Dordrecht, 1982); pp. 171—211.

[17] R. H. Möhring, Algorithmic aspects of comparability graphs and interval orders, in: *Graphs and Order* (I. Rival, ed.), D. Reidel (Dordrecht, 1985); pp. 41—101.

[18] O. Ore, *Theory of Graphs,* Amer. Math. Soc. (Providence, 1962).

[19] O. Ore, The four color problem, *Academic Press* (New York, 1967).

[20] C. R. Platt, Planar lattices and planar graphs, *J. Comb. Th. Ser B.,* **21** (1976), 30—39.

[21] W. R. Pulleyblank, On minimizing setups in precedence constrained scheduling, *Discrete Applied Math.,* to appear.

[22] I. Rival, Optimal linear extensions by interchanging chains, *Proc. Amer. Math. Soc.,* **89** (1983), 387—394.

[23] I. Rival, The diagram, in: *Graphs and Order* (I. Rival, ed.), D. Reidel (Dordrecht, 1985); pp. 103—133.

[24] L. N. Ševrin and N. D. Filippov, Partially ordered sets and their comparability graphs, *Sibirsk. Mat. Ž.,* **11** (1970), 648—667. (Russian)

(R. J.)
ÉCOLE DES MINES
SAINT-ÉTIENNE, FRANCE

(I. R.)
UNIVERSITY OF OTTAWA
DEPARTMENT OF COMPUTER SCIENCE
OTTAWA, K1N 9B4, CANADA

(R. N.)
DALHOUSIE UNIVERSITY
HALIFAX, CANADA

Acta Sci. Math., 51 (1987), 123—128

Representation of 2-distributive modular lattices of finite length

BJARNI JÓNSSON*) and J. B. NATION**)

Dedicated to the memory of Dr. András Huhn

1. Statement of the theorem. In this note we prove the following result.

Main Theorem. *Suppose L is a 2-distributive modular lattice of finite length n. If V is an n-dimensional vector space over a division ring D with $|D|>|L|$, then L can be embedded in $L(V)$.*

Here $L(V)$ is the lattice of all subspaces of the vector space V. To say that L is 2-distributive means that it satisfies the identity

$$a(x+y+z) = a(x+y)+a(x+z)+a(y+z).$$

A special case of this theorem, the case when L is of breadth 2, was proved in HERRMANN [3].

2. Preliminaries. A lattice is said to be n-distributive if it satisfies the identity

$$a \sum_{0\leq i\leq n} x_i = \sum_{0\leq i\leq n} a \sum_{j\neq i} x_j.$$

This concept was introduced by András Huhn and was investigated by him in a series of papers. His original definition, in HUHN [5], required the lattice to be modular, but this condition was dropped in [4]. We shall adhere to the revised terminology, although all the lattices under consideration here will be modular. The following two results will be needed.

Theorem A (HUHN [4], Theorem 3.1). *The dual of an n-distributive modular lattice is n-distributive.*

Received May 30, 1986.

*) Research supported by NSF Grant DMS 8300107

**) Research supported by NSF Grant MCS 83-01087

Theorem B (Huhn [5], Theorem 1.1). *An algebraic modular lattice is n-distributive iff it does not contain as a sublattice (the lattice of all subspaces of) a non-degenerate projective n-space.*

In one direction, this result can be strengthened: If an algebraic modular lattice fails to be n-distributive, then it contains a non-degenerate projective n-space as an interval, not just as a sublattice. This shows that the result proved here is in a sense "best possible". For suppose that L is a modular lattice of finite length, and that L is not 2-distributive. Then L contains as an interval a non-degenerate projective plane P. If P is not Arguesian, then there does not exist any embedding $L \rightarrowtail L(V)$, with V a vector space over a division ring D, but if P is Arguesian, then such embeddings can at most exist for division rings having the same characteristic as the coordinate ring of P.

In order not to have to interrupt the argument later, we state and prove here two simple observations that will be used in the proof of the main theorem.

Lemma C. *Suppose V is a finite dimensional vector space over a division ring D. Then V is not the union of fewer than $|D|$ proper subspaces of V.*

Proof. Suppose $|I|<|D|$, and suppose V_i, $i\in I$, are proper subspaces of V. Let $K=\bigcup_{i\in I} V_i$. Assuming that A is a subspace of V, we prove by induction on the dimension of A that if $A\subseteq K$, then $A\subseteq V_i$ for some $i\in I$. This is certainly true when $\dim A\leqq 1$. Assuming that $n>1$, and that the claim holds whenever $\dim A<n$, we consider the case $\dim A=n$. Then each proper subspace of A is contained in some V_i. Since A has $|D|$ subspaces of dimension $n-1$, it follows that at least two of these, say B and C, are contained in the same V_i, whence $A=B+C\subseteq V_i$.

Lemma D. *Suppose L and L' are modular lattices of the same finite lenght. If the mapping $f\colon L\to L'$ is one-to-one and preserves the covering relation, then f is an embedding of L into L'.*

Proof. By duality, it suffices to show that

$$f(a)f(b)=f(ab) \tag{1}$$

for all $a, b\in L$. The mapping is obviously monotone, so (1) holds whenever a and b are comparable. For the case when a and b are not comparable, we use induction on the lengths of the intervals a/ab and b/ab. If both intervals have length 1, then a and b are distinct covers of ab, and $f(a)$ and $f(b)$ are therefore distinct covers of $f(ab)$, so that (1) holds in this case. For the inductive step, we assume that $ab\prec c<a$ and let $d=b+c$. Then $ad=c$ and $bc=ab$, hence by the inductive hypothesis, $f(a)f(d)=f(c)$ and $f(b)f(c)=f(ab)$. Again using the monotonicity of f, we infer that (1) holds.

3. Proof of the Main Theorem. We are going to show that every embedding $F\colon a/0 \rightarrowtail A/0$, where a is a coatom in L and A is a coatom in $L(V)$, can be extended to an embedding $G\colon L \rightarrowtail L(V)$. From this the theorem follows by induction on n.

Let M be the set of all minimal elements of the set $\{x \in L\colon x \not\leqq a\}$. We need to look at some properties of the elements of M.

The elements of M are obviously join irreducible. For any $x \not\leqq a$, the element $a+x=1$ covers a, and ax is therefore covered by x. Consequently,

$$x = ax+p \quad \text{whenever} \quad x \geqq p \in M.$$

In particular,

$$p+q = p+a(p+q) \quad \text{for all} \quad p, q \in M.$$

The most important fact about the elements $p \in M$ is that the set

$$C_p = \{p+q\colon q \in M\}$$

is always a chain. In other words, any two joins $p+q$ and $p+r$, with $q, r \in M$, are comparable. Suppose this fails. Then $r \not\leqq p+q$, hence $r(p+q) \leqq a$, and similarly $q(p+r) \leqq a$. Consequently,

$$(p+q)(p+r)(q+r) = q(p+r)+r(p+q) \leqq a.$$

But using the dual of the 2-distributive law, we find that

$$\begin{aligned} &a+(p+q)(p+r)(q+r) = \\ &= [a+(p+q)(p+r)][a+(p+q)(q+r)][a+(p+r)(q+r)] \geqq \\ &\geqq (a+p)(a+q)(a+r) = 1. \end{aligned}$$

This contradiction proves that C_p is a chain.

Finally, we note that, for $p, q \in M$,

$$C_p \cap C_q = C_p \cap (1/(p+q)).$$

Certainly, the set on the left is included in the set on the right. To prove the opposite inclusion, consider any $c \in C_p \cap (1/(p+q))$. Then $c \geqq p+q$, and $c = p+r$ for some $r \in M$. Consequently $c = (q+p)+(q+r)$, and recalling that C_p is a chain, we infer that $c = q+p$ or $c = q+r$. In either case, $c \in C_q$, as was to be shown.

It will be convenient to have a fixed notation for the elements of the chain C_p, say

$$c_{p0} < c_{p1} < \ldots < c_{p\lambda_p}.$$

We also fix a one-to-one mapping $f\colon L \to D \setminus \{0\}$ and a vector $\xi \in V \setminus A$, and for each $c \in a/0$ we pick a vector $\alpha(c) \in F(c) \setminus \bigcup_{d<c} F(d)$. Such a vector always exists by Lemma C. Associating with each $p \in M$ the vector

$$\xi_p = \xi - \sum_{0 \leqq i < \lambda_p} f(c_{pi})\alpha(ac_{p(i+1)}),$$

we are now ready to describe the mapping $G: L \to L(V)$. For $x \leqq a$, we let $G(x) = F(x)$, but if $x \nleqq a$, then there exists $p \in M$ with $p \leqq x$. For each such p we let

$$G_p(x) = G(ax) + D\xi_p.$$

We claim that $G_p(x)$ is actually independent of p. To see this, consider another element $q \in M$ with $q \leqq x$. The element $p+q$ is in both C_p and C_q, say $p+q = c_{pk} = c_{qm}$. For $i<k$ we have $ac_{p(i+1)} \leqq a(p+q) \leqq ax$, and therefore $\alpha(ac_{p(i+1)}) \in G(ax)$. Consequently,

$$G_p(x) = G(ax) + D\big(\xi - \sum_{k \leqq i < \lambda_p} f(c_{pi})\alpha(ac_{p(i+1)})\big).$$

From this, and the corresponding formula for $G_q(x)$, we infer that $G_p(x) = G_q(x)$. Dropping the subscript, we therefore have a well defined mapping from L to $L(V)$.

It is easy to check that $x \leqq y$ implies $G(x) \leqq G(y)$. We need to check that $x<y$ implies $G(x) < G(y)$. Since $\xi \notin A$, this is clear when $x \leqq a$. If $x \nleqq a$, then $y = ay + x$, which implies that $ay \nleqq x$, so that $ax < ay$. But then $G(ay) \nleqq G(ax) + D\eta$ whenever $\eta \notin A$. Hence $G(x) < G(y)$.

The mapping G preserves strict inclusion, and since the lattices L and $L(V)$ have the same length, it follows that G preserves the covering relation. To prove that G is an embedding, it suffices by Lemma D to show that G is one-to-one. We shall do this by showing that $G(x) \leqq G(y)$ implies $x \leqq y$.

Suppose $G(x) \leqq G(y)$. If $x \leqq a$ and $y \leqq a$, then it obviously follows that $x \leqq y$. The case $x \nleqq a$ and $y \leqq a$ is excluded, for we would then have $G(x) \nleqq A$ and $G(y) \leqq A$. Next suppose $x \leqq a$ and $y \nleqq a$. Choosing $q \in M$ with $q \leqq y$, we then have $G(x) \leqq G(ay) + D\xi_q$, hence $G(x) \leqq A \cap (G(ay) + D\xi_q) = G(ay) + (A \cap D\xi_q) = G(ay)$; and consequently $x \leqq ay \leqq y$. Finally suppose $x \nleqq a$ and $y \nleqq a$. If $xy \nleqq a$, then we can choose $p \in M$ with $p \leqq xy$, and we have $x = ax + p$ and $y = ay + p$. From the fact that $G(ax) \leqq G(y)$, we infer by the preceding case that $ax \leqq y$, and since $p \leqq y$, we conclude that $x \leqq y$.

To complete the proof, it suffices to show that it cannot happen that $x \nleqq a$, $y \nleqq a$ and $xy \leqq a$. Assuming that these three conditions are satisfied, we choose $p, q \in M$ with $p \leqq x$ and $q \leqq y$. Then $G(x) = G(ax) + D\xi_p$ and $G(y) = G(ay) + D\xi_q$, and the condition $G(x) \leqq G(y)$ therefore implies that $\xi_p \in G(ay) + D\xi_q$. Thus $\xi_p = \eta + s\xi_q$ for some $\eta \in G(ay)$ and $s \in D$. Actually $s = 1$, because $\xi_p - s\xi_q \in A$. Let c_{pk} be the term in the chain C_p that precedes $p+q$, and let c_{qm} be the term in C_q that precedes $p+q$. Let

$$\beta_p = \sum_{0 \leqq i < k} f(c_{pi})\alpha(ac_{p(i+1)}), \quad \gamma_p = \sum_{k < i < \lambda_p} f(c_{pi})\alpha(ac_{p(i+1)}),$$

and define β_q and γ_q similarly. Then $\gamma_p = \gamma_q$, and therefore

$$\big(f(c_{pk}) - f(c_{qm})\big)\alpha\big(a(p+q)\big) = -\eta - \beta_p + \beta_q \in G(ay + ac_{pk} + ac_{qm}).$$

Since $\alpha(a(p+q))$ does not belong to $G(d)$ for any $d<a(p+q)$, it follows that

$$a(p+q) \leqq ay+ac_{pk}+ac_{qm}.$$

Recalling that $p+q=q+a(p+q)$, we infer that

$$p \leqq q+ay+ac_{pk}+ac_{qm}.$$

Since p is join irreducible and L is 2-distributive, it follows that p is included in the join of some two of the four elements q, ay, ac_{pk} and ac_{qm}. In fact, since $p \nleqq a$, we have $p \leqq q+w$ for some $w \in \{ay, ac_{pk}, ac_{qm}\}$. But each of these three inclusions readily leads to a contradiction: $p \leqq q+ay$ would imply $p \leqq y$, hence $xy \nleqq a$; $p \leqq q+ac_{pk}$ would imply $p \leqq (q+ac_{pk})c_{pk}=qc_{pk}+ac_{pk} \leqq a$; and $p \leqq q+ac_{qm}$ would imply $p \leqq c_{qm}$. This establishes the contradiction, and completes the proof of the theorem.

4. Connections with other representation problems. Modular lattices arise naturally in many contexts, and each source gives rise to a representation problem. The "most general" representation problem to receive extensive attention is the problem of representing a lattice as a lattice of permutable equivalence relations. The (modular) lattices for which such a representation exists are said to be of type 1. The discovery that there exist modular lattices that are not of type 1 led to the introduction of a six-variable identity, stronger than the modular law, that holds in every lattice of type 1. This identity holds in (the lattice of all subspaces of) a projective plane iff the plane is Arguesian, and the lattices in which the identity holds are therefore called Arguesian. Of course every modular lattice that contains a non-Arguesian projective plane as a sublattice is non-Arguesian, but as might be expected, the Arguesian identity can fail for other, more subtle reasons. However, the geometric flavor of these original examples carries over to a surprising extent to the general case. It is shown in Day and Jónsson [1] that if a modular lattice L fails to be Arguesian, then the ideal lattice of L contains a "non-Arguesian configuration" of ten points and ten lines. These twenty elements, however, may lie in up to twenty distinct non-degenerate planes that constitute intervals in the ideal lattice. In particular, therefore, every 2-distributive modular lattice L is Arguesian, for the ideal lattice of L is also 2-distributive and therefore does not contain a non-degenerate projective plane as a sublattice. We do not know whether every 2-distributive modular lattice is of type 1, although it seems likely that this is the case. In fact, we conjecture that a modified version of the result proved here is true without any restriction on the length of the lattice.

Conjecture. For any 2-distributive modular lattice L, and division ring D, L can be embedded in $L(V)$ for some vector space V over D.

After the research reported here was completed, we received a prepublication copy of a research announcement, HAIMAN [2], describing the construction of a lattice that is Arguesian but not of type 1. This important result confirms a conjecture of long standing, and it gives increased importance to various ongoing efforts to obtain positive representation results for special classes of modular lattices. Our result falls into that category, but it may have a special significance in this context. In DAY and JÓNSSON [1] it was shown that if a modular lattice fails to be Arguesian, then its ideal lattice either contains as an interval a bad plane, or else it contains two or more planes that are badly fitted together. It seems likely that a similar result holds for type 1. Our result gives some credence to this conjecture, for it shows that if a modular lattice of finite length is not of type 1, then it must contain as an interval a non-degenerate plane.

Bibliography

[1] A. DAY and B. JÓNSSON, A structural characterization of non-Arguesian lattices, *Order*, **2** (1986), 335—350.

[2] M. HAIMAN, Arguesian lattices which are not linear, *Bull. Amer. Math. Soc.*, **16** (1987), 121—124.

[3] C. HERRMANN, Quasiplanare Verbände, *Arch. Math. (Basel)*, **24** (1973), 240—246.

[4] A. P. HUHN, n-distributivity and some questions of the equational theory of lattices, in: *Contributions to Universal Algebra* (Proc. Conf. Szeged, 1975), Colloq. Math. Soc. J. Bolyai, vol. 17, North-Holland (Amsterdam, 1977); pp. 167—178.

[5] A. P. HUHN, Schwach distributive Verbände. I, *Acta Sci. Math.*, **33** (1972), 297—305.

[6] A. P. HUHN, Two notes on n-distributive lattices, in: *Lattice Theory* (Proc. Conf. Szeged, 1974), Colloq. Math. Soc. J. Bolyai, vol. 14, North-Holland (Amsterdam, 1976); pp. 137—147.

(B. J.)
DEPARTMENT OF MATHEMATICS
BOX 1543, STATION B
VANDERBILT UNIVERSITY
NASHVILLE, TENNESSEE 37235, U.S.A.

(J. B. N.)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF HAWAII AT MANOA
HONOLULU, HAWAII 96822, U.S.A.

Acta Sci. Math., 51 (1987), 129—135

Congruence relations and direct decompositions of ordered sets

M. KOLIBIAR

Dedicated to the memory of András Huhn

1. Introduction

Given an algebra $\mathscr{A}=(A; F)$, there is a one-one correspondence between direct product decompositions $\mathscr{A}=\prod(\mathscr{A}_i \mid i\in I)$ and families $(\theta_i \mid i\in I)$ of congruence relations of $\mathscr{A}$ satisfying

(1) $\bigcap(\theta_i \mid i\in I) = \mathrm{id}_A$,

(2) $\bigvee(\theta_i \mid i\in I) = A\times A$,

(3) given a family $(x_i \mid i\in I)$ of elements of A, there exists an element $x\in A$ such that $x\theta_i x_i$ for all $i\in I$.

The situation is more complicated in the case of relational systems. A method of characterization of direct product decompositions of such systems was given in the papers [1] and [3]. For the sake of simplicity we state the result for the case of ordered sets. (We use the term "ordered set" for partially ordered set.)

There is a one-one correspondence between direct product decompositions of an ordered set $\mathscr{A}=(A; \leqq)$ and families $(\theta_i \mid i\in I)$ of equivalence relations of A satisfying (1), (2), (3) and

(4) given elements x, y, x_i, y_i $(i\in I)$ of A such that $x_i\leqq y_i$ and $x\theta_i x_i, y\theta_i y_i$ for all $i\in I$, then $x\leqq y$.

The condition (4) is a kind of "collective congruence property". Instead of a notion of an (individual) congruence relation we have to deal with a "congruence family". Recently an analogous characterization of subdirect decompositions of multialgebras was given by G. E. HANSOUL [2].

In the present note we study a notion of congruence relation in the class of ordered sets which enables the same characterization of direct decompositions of

Received June 27, 1986.

directed ordered sets as in the case of algebras (conditions (1), (2), (3)) whenever the number of the decomposition factors is finite or the ordered set satisfies a chain condition. Moreover, we show that the congruence relations form a distributive lattice. Also some results on subdirect decompositions are given.

2. Congruence relations

2.1. Definition. Let $\mathcal{P}=(P;\leqq)$ be a directed ordered set. An equivalence relation θ on P will be called a congruence relation on $\mathcal{P}$ if the following conditions are satisfied.

(i) For each $a\in P$, $[a]\theta$ $(=\{x\in P \mid x\theta a\})$ is a convex subset of P.

(ii) If $a, b, c\in P$, $a\leqq c$, $b\leqq c$ and $a\theta b$, then there is $d\in P$ such that $a\leqq d\leqq c$, $b\leqq d$ and $a\theta d$.

(iii) If $a, b, u, v\in P$, $u\leqq a\leqq v$, $u\leqq b\leqq v$ and $u\theta a$ ($a\theta v$), then there is $t\in P$ such that $b\leqq t\leqq v$, $a\leqq t$, ($u\leqq t\leqq b$, $t\leqq a$) and $b\theta t$.

If $\mathcal{P}$ is a lattice then this notion coincides with the lattice congruence relation.

2.2. It can be easily shown that the conditions (ii), (iii) are equivalent with the following one:

(iv) If $a, b, c, d\in P$, $a\leqq c\leqq d$, $b\leqq d$ ($a\geqq c\geqq d$, $b\geqq d$) and $a\theta b$, then there is $e\in P$ such that $c\leqq e\leqq d$, $b\leqq e$ ($c\geqq e\geqq d$, $b\geqq e$) and $c\theta e$.

2.3. Let Con $\mathcal{P}$ denote the set of all congruence relations of $\mathcal{P}$. Eq P will denote the lattice of all equivalence relations on P.

In what follows all ordered sets will be supposed to be (both up- and down-) directed (i.e. to any a, b there are u, v such that $u\leqq a\leqq v$, $u\leqq b\leqq v$). $\mathcal{P}$ will denote an ordered set $(P;\leqq)$. We say that $\mathcal{P}$ satisfies the restricted ascending chain condition (RACC) if every closed interval of $\mathcal{P}$ satisfies the ascending chain condition.

The set $\{1, 2, ..., n\}$ will be denoted by $\bar{n}$.

2.4. *A congruence relation θ on $\mathcal{P}$ has the following property.*

(ii') *If $a, b, c\in P$, $a\geqq c$, $b\geqq c$, $a\theta b$, then there exists $e\in P$ such that $a\geqq e\geqq c$, $b\geqq e$ and $a\theta e$.*

Proof. Using the fact that $\mathcal{P}$ is up-directed and (ii) we get that $d\in P$ exists such that $a\leqq d$, $b\leqq d$ and $b\theta d$. The existence of the desired element e follows by (iii).

2.5. *Let $\theta\in$ Con $\mathcal{P}$, $a, b\in P$, $a\leqq b$. Then to any $x\in[a]\theta$ ($y\in[b]\theta$) there is $y\in[b]\theta$ ($x\in[a]\theta$) such that $x\leqq y$.*

Proof. Let $x \in [a]\theta$. Since $\mathcal{P}$ is up-directed, there is $c \in P$ such that $b \leqq c$, $x \leqq c$. According to 2.2 there is $y \in P$ such that $b\theta y$ and $x \leqq y$. The second assertion is symmetric.

2.6. *Let* $\theta \in \mathrm{Con}\,\mathcal{P}$. *Given* $a, b \in P$, *set* $[a]\theta \leqq [b]\theta$ *if there are* $x, y \in P$ *such that* $x\theta a$, $y\theta b$ *and* $x \leqq y$. *Then* $(P/\Theta; \leqq)$ *is an ordered set.*

The proof is straightforward.

2.7. *Let* $\theta \in \mathrm{Con}\,\mathcal{P}$, $a, b, c, d, u \in P$, $a \leqq b$, $a\theta c$, $b\theta d$, $c \leqq u$, $d \leqq u$. *Then there is* $v \in P$ *such that* $d \leqq v \leqq u$, $c \leqq v$ *and* $b\theta v$.

Proof. According to 2.5 there exists $e \in P$ such that $e\theta a$ and $e \leqq d$. Using 2.2 we get that $v \in P$ exists such that $d \leqq v \leqq u$, $c \leqq v$ and $d\theta v$.

2.8. Theorem 1. *The congruence relations on* $\mathcal{P}$ *correspond one-one to mappings* f *of* P *onto ordered sets (given uniquely up to isomorphism) such that*

(a) f *is isotone,*

(b) *if* $x, y, u \in P$, $x \leqq u$, $y \leqq u$ $(x \geqq u,\ y \geqq u)$ *and* $f(x) \leqq f(y)$ $(f(x) \geqq f(y))$ *then there is* $z \in P$ *such that* $x \leqq z$, $y \leqq z \leqq u$ $(x \geqq z,\ y \geqq z \geqq u)$ *and* $f(y) = f(z)$.

Proof. Let $\theta \in \mathrm{Con}\,\mathcal{P}$. The mapping f: $P \to P/\theta$, $x \mapsto [x]\theta$, is obviously isotone. Let $f(x) \leqq f(y)$ and $x \leqq u$, $y \leqq u$. Then $a \in [x]\theta$ and $b \in [y]\theta$ exist such that $a \leqq b$. The existence of the desired element z follows by 2.7. The second assertion follows by symmetry.

Conversely, let f: $P \to Q$ fulfil the conditions (a) and (b) and let $\theta = \mathrm{Ker}\, f$. The property (i) is obvious. Let $a\theta b$ and $a \leqq c$, $b \leqq c$. Then $f(a) = f(b)$ and according to (b) there is $d \in P$ sucht that $a \leqq d \leqq c$, $b \leqq d$ and $f(a) = f(d)$ which proves (ii). If $u \leqq a \leqq v$, $u \leqq b \leqq v$ and $u\theta a$, then $f(a) \leqq f(b)$ and, according to (b), $t \in P$ exists such that $b \leqq t \leqq v$, $a \leqq t$ and $f(t) = f(b)$, hence $t\theta b$. The second part of (iii) follows by symmetry.

Let F and G be the mappings $\theta \mapsto f$ and $f \mapsto \theta$, respectively, described above. Obviously $(GF)(\theta) = \theta$. Let f: $P \to Q$ be given. Then the mapping $(FG)(f)$ is the canonical mapping $P \to P/\theta$, $\theta = \mathrm{Ker}\, f$. To prove $\mathcal{Q} \cong \mathcal{P}/\theta$ let us define h: $P/\theta \to Q$ by setting $h([x]\theta) = f(x)$. h is surjective and well defined. If $[x]\theta \leqq [y]\theta$ then $z \in [y]\theta$ exists such that $x \leqq z$. Hence $f(x) \leqq f(z) = f(y)$. Conversely, let $f(x) \leqq f(y)$. There is $u \in P$ such that $x \leqq u$, $y \leqq u$. According to (b) $z \in P$ exists such that $y \leqq z \leqq u$, $x \leqq z$, $f(y) = f(z)$, hence $[x]\theta \leqq [z]\theta = [y]\theta$.

2.9. *Let* $x_0 \leqq x_1 \leqq \ldots \leqq x_n$, $x_0 \leqq y \leqq x_n$, $\alpha_i \in \mathrm{Con}\,\mathcal{P}$, $x_{i-1}\alpha_i x_i$ *for all* $i \in \bar{n}$. *Then there exists a sequence* $x_0 = y_0 \leqq y_1 \leqq \ldots \leqq y_n = y$ *such that* $y_i \leqq x_i$ *and* $y_{i-1}\alpha_i y_i$ *for all* $i \in \bar{n}$.

Proof. If $n=1$, it suffices to take $y_1 = y$. Suppose the assertion holds for $n-1 \geqq 1$. Using (iii) for the elements y, x_{n-1}, x_0, x_n we get that y_{n-1} exists such that $x_0 \leqq y_{n-1} \leqq y$, $y_{n-1} \leqq x_{n-1}$ and $y_{n-1} \alpha_n y$. By the induction assumption there are elements $y_1, \ldots, y_{n-2}$ which together with y_{n-1} give the desired sequence.

2.10. *Let $u \leqq a \leqq v$, $u \leqq b \leqq v$, $u = u_0 \leqq u_1 \leqq \ldots \leqq u_n = a$, $u_{i-1} \alpha_i u_i$, $\alpha_i \in \mathrm{Con}\, \mathcal{P}$ for all $i \in \bar{n}$. Then a sequence $b = v_0 \leqq v_1 \leqq \ldots \leqq v_n \leqq v$ exists such that $a \leqq v_n$, $u_i \leqq v_i$ and $v_{i-1} \alpha_i v_i$ for all $i \in \bar{n}$.*

Proof. If $n=1$, the assertion follows by (iii). Assume the assertion holds for $n-1$. By (iii) there is v_1 such that $u_1 \leqq v_1 \leqq v$, $b \leqq v_1$ and $b \alpha_1 v_1$. By the induction assumption there exists a sequence $v_1 \leqq v_2 \leqq \ldots \leqq v_n \leqq v$ such that $a \leqq v_n$, $u_i \leqq v_i$ and $v_{i-1} \alpha_i v_i$ for $i = 2, \ldots, n$.

2.11. *Let $a = t_0, t_1, \ldots, t_n = b$ be elements of P, $\alpha_i \in \mathrm{Con}\, \mathcal{P}$ and $t_{i-1} \alpha_i t_i$ for all $i \in \bar{n}$. Then there exist sequences $a = u_0 \leqq u_1 \leqq \ldots \leqq u_n$, $b = v_0 \leqq v_1 \leqq \ldots \leqq v_n = u_n$ such that for each $i \in \bar{n}$, $u_{i-1} \alpha_{j(i)} u_i$, $v_{i-1} \alpha_{k(i)} v_i$ where $j(i), k(i) \in \bar{n}$.*

Proof. In $n=1$, the assertion is trivial. Supposing the assertion holds for $n-1$ we shall prove it for n. Using the induction assumption for the elements $a, t_1, \ldots, t_{n-1}$ we get sequences $a = u_0 \leqq u_1 \leqq \ldots \leqq u_{n-1}$, $t_{n-1} = w_0 \leqq w_1 \leqq \ldots \leqq w_{n-1} = u_{n-1}$ such that, for each $i \in \overline{n-1}$, $u_{i-1} \alpha_{j(i)} u_i$, $w_{i-1} \alpha_{k(i)} w_i$, $j(i), k(i) \in \overline{n-1}$. Using the fact that P is up-directed and (ii), (iii), we get that $c, v_1 \in P$ exist such that $u_{n-1} \leqq c$, $t_{n-1} \leqq v_1$, $b \leqq v_1 \leqq c$, $t_{n-1} \alpha_n v_1 \alpha_n b$ and $u_{n-1} \alpha_n c$. According to 2.10 there exists a sequence $v_1 \leqq v_2 \leqq \ldots \leqq v_n \leqq c$ such that ($w_{i-1} \leqq v_i$ for $i \in \bar{n}$ and) $u_{n-1} \leqq v_n$, $v_{i-1} \alpha_{k(i)} v_i$ for $i = 2, \ldots, n$. Obviously $u_{n-1} \alpha_n u_n$, where $u_n = v_n$.

2.12. *Let $A \subset \mathrm{Con}\, \mathcal{P}$. Then $\bigvee(\alpha \mid \alpha \in A) = \beta$ has the property* (ii).

Proof. Let $a \leqq c$, $b \leqq c$, $a \beta b$. According to the proposition dual to 2.11 there exists a sequence $u_0 \leqq u_1 \leqq \ldots \leqq u_n = b$ such that $u_0 \leqq a$ and $u_{i-1} \alpha_i u_i$, $\alpha_i \in A$ for each $i \in \bar{n}$. According to 2.10 there exists a sequence $a = v_0 \leqq v_1 \leqq \ldots \leqq v_n \leqq c$ such that $b \leqq v_n$ and $v_{i-1} \alpha_i v_i$ for all $i \in \bar{n}$. Hence $a \beta v_n$ (and $b \beta v_n$).

2.13. *Let $A \subset \mathrm{Con}\, \mathcal{P}$, $(x, y) \in \bigvee(\alpha \mid \alpha \in A)$ and $x \leqq y$. Then there exists a sequence $x = x_0 \leqq x_1 \leqq \ldots \leqq x_n = y$ such that $x_{i-1} \alpha_i x_i$, $\alpha_i \in A$ for all $i \in \bar{n}$.*

Proof. According to 2.11 there exists a sequence $x = u_0 \leqq u_1 \leqq \ldots \leqq u_n$ such that $y \leqq u_n$ and $u_{i-1} \alpha_{j(i)} u_i$, $\alpha_{j(i)} \in A$ for all $i \in \bar{n}$. According to 2.9 there exists a sequence $x = t_0 \leqq t_1 \leqq \ldots \leqq t_n = y$ such that $t_{i-1} \alpha_{j(i)} t_i$ for all $i \in \bar{n}$.

2.14. *If $A \subset \mathrm{Con}\, \mathcal{P}$ then $\beta = \bigvee(\alpha \mid \alpha \in A)$ has the property* (i).

Proof. Let $x \leqq z \leqq y$ and $x \theta y$. Using 2.13 and 2.9 we get $x \theta z$.

2.15. *If $A \subset \mathrm{Con}\, \mathcal{P}$ then $\bigvee(\alpha \mid \alpha \in A) \in \mathrm{Con}\, \mathcal{P}$.*

Proof. The property (iii) of the join follows immediately from 2.13 and 2.10 (and from its duals), while (i) and (ii) were proved in 2.14 and 2.12.

2.16. *If* $\alpha, \beta \in \mathrm{Con}\,\mathscr{P}$ *then* $\alpha \cap \beta \in \mathrm{Con}\,\mathscr{P}$.

Proof. Obviously $\alpha \cap \beta$ has the property (i). The properties (ii) and (iii) can be easily checked.

2.17. From 2.15 and 2.16 we get

Theorem 2. Con $\mathscr{P}$ *forms a complete lattice which is a sublattice of the lattice* Eq P. *Moreover for any set* $A \subset \mathrm{Con}\,\mathscr{P}$, $\bigvee_{\mathrm{Con}\,\mathscr{P}}(\alpha \mid \alpha \in A) = \bigvee_{\mathrm{Eq}\,P}(\alpha \mid \alpha \in A)$.

Remark. Unfortunately, the set-theoretic intersection $\bigcap(\alpha \mid \alpha \in A)$ need not belong to Con $\mathscr{P}$ if A is an infinite subset of Con $\mathscr{P}$, as the following simple example shows. Let N be the set of all negative integers with the natural order and $P = \{u, a, b\} \cup N$, $u < a < n$ and $u < b < n$ for all $n \in N$. For each $n \in N$, let α_n be the equivalence relation on P with the blocks $(n]$ and $[n+1)$ ($\emptyset$ if $n = -1$). Then $\alpha_n \in \mathrm{Con}\,\mathscr{P}$ but $\bigcap(\alpha_n \mid n \in N) \notin \mathrm{Con}\,\mathscr{P}$.

Hence it can occur that $\bigwedge_{\mathrm{Con}\,\mathscr{P}}(\alpha \mid \alpha \in A) < \bigwedge_{\mathrm{Eq}\,P}(\alpha \mid \alpha \in A)$ if A is infinite.

Theorem 3. *The lattice* Con $\mathscr{P}$ *is distributive. If* $\alpha \in \mathrm{Con}\,\mathscr{P}$ *and* $B \subset \mathrm{Con}\,\mathscr{P}$ *then* $\alpha \wedge \bigvee(\beta \mid \beta \in B) = \bigvee(\alpha \wedge \beta \mid \beta \in B)$.

Proof. Set $\varphi = \alpha \wedge \bigvee(\beta \mid \beta \in B)$, $\psi = \bigvee(\alpha \wedge \beta \mid \beta \in B)$. Obviously $\psi \leqq \varphi$. To prove the converse we first notice that $x\varphi y$ and $x \leqq y$ imply $x\psi y$. Indeed, from the assumption we get $x\alpha y$ and the existence of a sequence $x = x_0 \leqq x_1 \leqq \ldots \leqq x_n = y$, $x_{i-1}\beta_i x_i$, $\beta_i \in B$, for all $i \in \bar{n}$ (2.13). Then $x_{i-1}\ \alpha \wedge \beta_i\ x_i$ hence $x\psi y$. To get the implication for arbitrary $x, y \in P$, observe that if $x\varphi y$ then $z \in P$ exists such that $x \leqq z$, $y \leqq z$ and $x\varphi z$, $y\varphi z$.

2.18. *Let* $\mathscr{P}$ *satisfy RACC and let* $\alpha_i \in \mathrm{Con}\,\mathscr{P}$ $(i \in I)$, $\bigcap(\alpha_i \mid i \in I) = \mathrm{id}_P$. *If* a *and* a_i $(i \in I)$ *are elements of* P *such that* $a \leqq a_i$ *and* $a\,\alpha_i a_i$ *for all* $i \in I$, *then* $a = \inf(a_i \mid i \in I)$.

Proof. Let $b \leqq a_i$ for all $i \in I$. Choose $i(1) \in I$. According to (iii) there is $b_1 \in P$ such that $b_1 \leqq a$, $b_1 \leqq b$ and $b\alpha_{i(1)} b_1$. Choose $i(2) \in I - \{i(1)\}$. Then there exists $b_2 \in P$ such that $b_1 \leqq b_2 \leqq a$, $b_2 \leqq b$ and $b\alpha_{i(2)} b_2$. By induction we get a sequence $b_1 \leqq b_2 \leqq \ldots$, $b_j \leqq a$, $b_j \leqq b$, $b\alpha_{i(j)} b_j$, which ends with some member b_m by virtue of RACC. Then $b\alpha_j b_m$ for each $j \in I$ hence $b = b_m$ so that $b \leqq a$.

By an analogous argument we get the following proposition.

Let $\mathscr{P}$ *be an arbitrary directed ordered set. If there are given* $a, a_1, \ldots, a_n \in P$ *and* $\alpha_i \in \mathrm{Con}\,\mathscr{P}$ *with* $\alpha_1 \wedge \ldots \wedge \alpha_n = \mathrm{id}_P$ *such that* $a \leqq a_i$ *and* $a\,\alpha_i a_i$ *for each* $i \in \bar{n}$ *then* $a = \inf(a_1, \ldots, a_n)$.

Remark. Without the condition RACC the first proposition would not be true as the following example shows.

Let P be the set $A \cup B \cup \{a, b\}$ where A (B) is the set of all positive (negative) integers with its natural order and for any $m \in A$ and $n \in B$ let $m < a < n$, $m < b < n$. Then $\mathcal{P} = (P; \leqq)$ is a directed ordered set. For each $m \in A$, let α_m be the equivalence relation on P in which the only non-singleton blocks are the intervals $[a, -m]$ and $[m, b]$. Then $\alpha_m \in \mathrm{Con}\,\mathcal{P}$, $\bigcap(\alpha_m \mid m \in A) = \mathrm{id}_P$, $a \leqq n$, $b \leqq n$ and $a\alpha_{-n}n$ for each $n \in B$ but $b \leqq a$ does not hold.

3. Direct and subdirect representations

3.1. Definition. A subdirect product $\mathcal{P} \to \prod(\mathcal{P}_i \mid i \in I)$ will be called a full subdirect product whenever to each $i \in I$ and any $a, b \in P$ there is $c \in P$ such that $c_i = a_i$ and $c_j = b_j$ for all $j \neq i$.

3.2. Theorem 4. *Let $\mathcal{P} \to \prod(\mathcal{P}_i \mid i \in I)$ be a full subdirect product of ordered sets and let, for each $i \in I$, θ_i be the kernel of the projection $\mathcal{P} \to \mathcal{P}_i$. Then $\theta_i \in \mathrm{Con}\,\mathcal{P}$.*

Proof. Obviously θ_i fulfils (i). Let $a \leqq c$, $b \leqq c$ and $a\theta_i b$, i.e., $a_i = b_i$. Let d be the element of P with $d_i = a_i$ and $d_j = c_j$ for $j \neq i$. Then d is the element needed for (ii). The dual part of (ii) is analogous. Finally, let the elements $a, b, u, v \in P$ satisfy $u \leqq a \leqq v$, $u \leqq b \leqq v$, $u\theta_i a$ and let d be the element fulfilling $d_i = b_i$ and $d_j = v_j$ for $j \neq i$. Then d dulfils the condition of (iii). The dual part is analogous.

Remark. The theorem would not be true if the word "full" was omitted. This is shown by the following example. Let $L = \{o, i, a, b, c\}$ be the five-element modular and non-distributive lattice and C the chain $0 < 1 < 2$. Then $f\colon L \to C \times C$, where $o \mapsto (0, 0)$, $a \mapsto (0, 2)$, $b \mapsto (1, 1)$, $c \mapsto (2, 0)$, $i \mapsto (2, 2)$, gives a subdirect decomposition of the ordered set L but the kernels of the corresponding projections do not fulfil condition (iii).

3.3. Theorem 5. *Let, for each $i \in \bar{n}$, $\alpha_i \in \mathrm{Con}\,\mathcal{P}$ and $\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n = \mathrm{id}_P$. Then $\mathcal{P}$ is a subdirect product of the ordered sets $\mathcal{P}/\alpha_i$, where $x \mapsto ([x]\alpha_i \mid i \in \bar{n})$.*

Proof. It suffices to show that $[x]\alpha_i \leqq [y]\alpha_i$ for all $i \in \bar{n}$ implies $x \leqq y$. For each $i \in \bar{n}$ there exists $y_i \in P$ such that $y \leqq y_i$, $y\alpha_i y_i$ and $x \leqq y_i$ (see 2.5). According to the second proposition in 2.18, $y = \inf(y_1, \ldots, y_n)$, hence $x \leqq y$.

By an analogous argument (using the first proposition 2.18) we get

Theorem 6. *Let $\mathcal{P}$ satisfy the RACC and let A be a subset of $\mathrm{Con}\,\mathcal{P}$ such that $\bigcap(\theta \mid \theta \in A) = \mathrm{id}_P$. Then $\mathcal{P}$ is a subdirect product of the ordered sets $\mathcal{P}/\theta$ $(\theta \in A)$, where $x \mapsto ([x]\theta \mid \theta \in A)$.*

Theorem 7. *Let $\mathcal{P}$ be a (directed) ordered set. There is a one-one correspondence between direct product decompositions of $\mathcal{P}$ into finitely many (say n) factors and the families $(\theta_i \mid i \in \bar{n})$ of congruence relations of $\mathcal{P}$ satisfying* (1), (2) *and* (3) (see the introduction).

Proof. The theorem is an easy consequence of Theorems 4 and 5.

Analogously (using Theorem 6) the following theorem can be proved.

Theorem 8. *Let $\mathcal{P}$ satisfy RACC.*

(a) *There is a one-one correspondence between the direct product decompositions $\mathcal{P} \to \prod(\mathcal{P}_i \mid i \in I)$ and the families $(\theta_i \mid i \in I)$ of congruence relations satisfying* (1), (2) *and* (3).

(b) *There is a one-one correspondence between the full subdirect product decompositions $\mathcal{P} \to \prod(\mathcal{P}_i \mid i \in I)$ and the families $(\theta_i \mid i \in I)$ of congruence relations satisfying* (1) *and*

(5) *for each $i \in I$, $\theta_i \circ \bigcap(\theta_j \mid j \in I,\ j \neq i) = P \times P$.*

Theorem 9. *Let $\mathcal{P}$ satisfy RACC and let $(\theta_i \mid i \in I)$ be a family of congruence relations of $\mathcal{P}$ satisfying* (1) *and* (5). *Then, for any subset $J \subset I$, $\bigcap(\theta_j \mid j \in J)$ (set-theoretic intersection) belongs to* Con $\mathcal{P}$.

Proof. According to Theorem 8 the family $(\theta_i \mid i \in I)$ gives a full subdirect product decomposition of $\mathcal{P}$. Then $\varphi = \bigcap(\theta_j \mid j \in J)$ and $\psi = \bigcap(\theta_k \mid k \in I - J)$ are equivalence relations corresponding to the direct product $\mathcal{P}/\varphi \times \mathcal{P}/\psi$, hence they belong to Con $\mathcal{P}$ (see Theorem 4).

Added in proof. 1. Recently J. Jakubík showed that the condition RACC in Theorem 8 cannot be omitted.

2. The condition (i) in 2.1 is an easy consequence of (iii) (this was observed by Mrs. J. Lihová). Also the condition (2) in the introduction may be omitted.

References

[1] C. C. Chang, B. Jónsson and A. Tarski, Refinement properties for relational structures, *Fund. Math.*, **55** (1964), 249—281.

[2] G. E. Hansoul, A subdirect decomposition theorem for multialgebras, *Algebra Universalis*, **16** (1983), 275—281.

[3] M. Kolibiar, On direct products of relational systems, *Acta Fac. Nat, Univ, Comenian, Math,.* **XII** (1965), 1—9. (German)

KATEDRA ALGEBRY A TÉORIE ČÍSEL
UNIVERZITA KOMENSKÉHO
842 15 BRATISLAVA, CZECHOSLOVAKIA

Acta Sci. Math., 51 (1987), 137—144

On circuits of atoms in atomistic lattices

SHUICHIRO MAEDA

Dedicated to the memory of Dr. András P. Huhn

1. Introduction. In [2], an atomistic lattice with the covering property is called an AC-lattice and an upper continuous AC-lattice is called a matroid lattice. (In [1], [6] and [9], a matroid lattice is called a geometric lattice.) As shown in [9], Section 3.3, a finite lattice is a matroid lattice if and only if it is isomorphic to the lattice of closed sets of a matroid. By this reason, the concept of circuits which plays an important role in matroid theory can be introduced in the theory of matroid lattices. The main purpose of this paper is to investigate lattice-theoretical properties of circuits.

In Section 2, we shall show that the set of atoms of an atomistic join-semilattice L with the finite covering property forms a simple matroid. The set $F(L)$ of all finite elements of L forms an AC-lattice, which will be called an FAC-lattice, and in Section 3, we define circuits of atoms in an FAC-lattice.

In Section 4, we shall discuss a connection between the modularity of FAC-lattices and the existence of special circuits, which will be called P-circuits. An important example of FAC-lattice is a bond lattice associated with a non-oriented finite simple graph ([6], [8]). Such a lattice has a remarkable property, that is, it has no non-trivial P-circuit.

Another important example is an affine matroid lattice whose properties are thoroughly investigated in [2], Chapter IV. This lattice always has a property called strongly planar. In Section 5, we shall show that almost all non-modular bond lattices are not strongly planar. From this we can see that, in the set of matroid lattices, there are three disjoint subsets: {non-modular affine matroid lattices}, {non-modular bond lattices} and {modular matroid lattices}.

Received May 14, 1986.

2. Atomistic join-semilattices and simple matroids. Let L be a join-semilattice with 0. The set of all atoms of L is denoted by $\Omega(L)$. We put

$$F(L) = \{p_1\vee\dots\vee p_n;\ \ p_i\in\Omega(L),\ \ n=1,2,\dots\}\cup\{0\},$$

and an element of $F(L)$ is called a *finite element*. L is called *atomistic* when every non-zero element a of L is the least upper bound of $\{p\in\Omega(L);\, p\leqq a\}$ (see [7]).

Let L be an atomistic join-semilattice. For any subset ω of $\Omega(L)$, the *closure* of ω is defined by

$$\mathrm{Cl}(\omega)=\{p\in\Omega(L);\ p\leqq q_1\vee\dots\vee q_n,\ q_i\in\omega\}\ (\mathrm{Cl}(\emptyset)=\emptyset).$$

The following properties are easily verified.

(Cl 1) $\omega\subset\mathrm{Cl}(\omega)$.

(Cl 2) $\omega_1\subset\omega_2$ implies $\mathrm{Cl}(\omega_1)\subset\mathrm{Cl}(\omega_2)$.

(Cl 3) $\mathrm{Cl}(\mathrm{Cl}(\omega))=\mathrm{Cl}(\omega)$.

(Cl 4) If $p\in\mathrm{Cl}(\omega)$ then there exists a finite subset ω' of ω such that $p\in\mathrm{Cl}(\omega')$.

(Cl 5) $\mathrm{Cl}(\{p\})=\{p\}$ for $p\in\Omega(L)$.

Proposition 1. *Let L be an atomistic join-semilattice.*

(i) *The following two statements are equivalent:*

(α) *L has the finite covering property, i.e., if $p\in\Omega(L)$, $a\in F(L)$ and $p\nleqq a$ then $a\vee p$ covers a.*

(β) *If $p, q\in\Omega(L)$, $p\in\mathrm{Cl}(\omega\cup\{q\})$ and $p\notin\mathrm{Cl}(\omega)$ then $q\in\mathrm{Cl}(\omega\cup\{p\})$.*

(ii) *If L satisfies (α) (and (β)), then $F(L)$ is an AC-lattice and $M(L)=(\Omega(L),\mathrm{Cl})$ is a simple matroid. Moreover, the set $L(M(L))=\{\omega\subset\Omega(L);\ \mathrm{Cl}(\omega)=\omega\}$ forms a matroid lattice by set-inclusion, and $F(L)$ is lattice-isomorphic to $F(L(M(L)))$ by the mapping*

$$a\mapsto\omega(a)=\{p\in\Omega(L);\ p\leqq a\}.$$

Proof. (i) It follows from [7], Theorem 2.2 that (α) is equivalent to the following statement (exchange property):

(α') If $p, q\in\Omega(L)$, $a\in F(L)$, $p\leqq a\vee q$ and $p\nleqq a$ then $q\leqq a\vee p$.

We shall prove (α')$\Rightarrow$(β). If $p\in\mathrm{Cl}(\omega\cup\{q\})$ and $p\notin\mathrm{Cl}(\omega)$, then there exist $r_1,\dots,r_n\in\omega$ such that $p\leqq r_1\vee\dots\vee r_n\vee q$ and $p\nleqq r_1\vee\dots\vee r_n$. Hence, $q\leqq r_1\vee\dots\vee r_n\vee p$ by (α'), and hence $q\in\mathrm{Cl}(\omega\cup\{p\})$. ($\beta$)$\Rightarrow$($\alpha'$). Let $a\in F(L)$, $p\leqq a\vee q$ and $p\nleqq a$. We put $a=r_1\vee\dots\vee r_n$, $r_i\in\Omega(L)$ and $\omega=\{r_1,\dots,r_n\}$. Then, $p\in\mathrm{Cl}(\omega\cup\{q\})$ and $p\notin\mathrm{Cl}(\omega)$. Hence, $q\in\mathrm{Cl}(\omega\cup\{p\})$ by (β), and hence $q\leqq a\vee p$.

(ii) If L satisfies (α), then $F(L)$ is a lattice by [7], Theorem 2.5. Evidently, $F(L)$ is atomistic and has the covering property. Moreover, $M(L)=(\Omega(L),\mathrm{Cl})$ is a matroid, since the closure operator satisfies (Cl 1)$\sim$(Cl 4) and (β) ([9], 1.2 and 20.2). $M(L)$ is simple by (Cl 5) and $\mathrm{Cl}(\emptyset)=\emptyset$ ([9], 1.4). The last statement follows from [2], (15.5) and (15.7), since ω is a subspace in the sense of [2], (15.1), if and only if $\mathrm{Cl}(\omega)=\omega$.

Definition. If an AC-lattice L satisfies $F(L)=L$, we shall call it an *FAC-lattice.*

The mapping $L\mapsto F(L)$ is a bijection between the set of matroid lattices and the set of FAC-lattices, because if L is a matroid lattice then $F(L)$ is an FAC-lattice and L is isomorphic to the lattice of all ideals of $F(L)$ by [2], (15.5) and (15.7).

Hereafter, we shall investigate properties of FAC-lattices. We remark the following facts (see [2], (8.5) and (8.14)). Each element a of an FAC-lattice L has the height $h(a)$, and $h(a)=n$ $(a\neq 0)$ if and only if there exist $p_1, \ldots, p_n\in\Omega(L)$ such that $a=p_1\vee\ldots\vee p_n$ and $(p_1\vee\ldots\vee p_{i-1})\wedge p_i=0$ for $i=2, \ldots, n$. For $a, b\in L$, we have

$$h(a\vee b)+h(a\wedge b)\leqq h(a)+h(b),$$

and equality holds if and only if (a, b) is a modular pair (denoted by $(a, b)M$).

3. Circuits of atoms. In this section, let L be an FAC-lattice.

Lemma 2. *Let $\omega=\{p_1, \ldots, p_n\}$ be a finite subset of $\Omega(L)$. The following statements are equivalent.*

(α) $(p_1\vee\ldots\vee p_{i-1})\wedge p_i=0$ *for* $i=2, \ldots, n$.

(β) *ω is a semi-orthogonal family, i.e., if ω_1, ω_2 are disjoint subsets of ω then $\bigvee(p;\ p\in\omega_1)\perp\bigvee(p;\ p\in\omega_2)$, where $a\perp b$ means $a\wedge b=0$ and $(a, b)M$* ([2], (2.2) and (8.12)).

(γ) *ω is an independent set of the matroid $M(L)=(\Omega(L), \mathrm{Cl})$, i.e., $p_i\notin\mathrm{Cl}(\omega-\{p_i\})$ for every i* (see [9], 1.7).

(δ) $h(p_1\vee\ldots\vee p_n)=n$.

Proof. (γ) is equivalent to the following statement: $p_i\wedge\bigvee_{j\neq i}p_j=0$ for every i. Hence, the implications (β)$\Rightarrow$(γ)$\Rightarrow$(α) are evident. (α)$\Rightarrow$(β) follows from [2], (2.5) and (8.12). Finally, the equivalence of (α) and (δ) follows from [2], (8.4).

Definition. As in matroid theory, we call a finite subset ω of $\Omega(L)$ a *circuit* when ω is a minimal dependent set, i.e., $\omega-\{p\}$ is independent and $p\leqq\bigvee(q;\ q\in\omega-\{p\})$ for any $p\in\omega$. For instance, if p, q, r are different atoms and $p\leqq q\vee r$ then $\{p, q, r\}$ is a circuit. The cardinality $|C|$ of a circuit C is not less than 3. The set of all circuits of $\Omega(L)$ is denoted by $\mathscr{C}(L)$.

Proposition 3. *Let a, b be elements of an FAC-lattice L. The following statements are equivalent.*

(α) *a and b are semi-orthogonal.*

(β) $h(a\vee b) = h(a)+h(b)$.

(γ) $\omega(a)\cap\omega(b)=\emptyset$ *and there is no* $C\in\mathscr{C}(L)$ *such that* $C\subset\omega(a)\cup\omega(b)$, $C\not\subset\omega(a)$ *and* $C\not\subset\omega(b)$.

Proof. (β)$\Rightarrow$(α). By (β) we have $h(a\vee b)+h(a\wedge b)\leqq h(a)+h(b)=h(a\vee b)$. Hence, $h(a\wedge b)=0$ and $(a, b)M$, so that $a\perp b$.

(α)$\Rightarrow$(γ). $\omega(a)\cap\omega(b)=\emptyset$ is evident. Let $C\in\mathscr{C}(L)$ with $C\subset\omega(a)\cup\omega(b)$. If both $\omega(a)\cap C$ and $\omega(b)\cap C$ were independent sets, then $C=(\omega(a)\cap C)\cup(\omega(b)\cap C)$ would be a semi-orthogonal family by (α) and [2], (2.4). This contradicts that C is a dependent set. Hence, for instance, $\omega(a)\cap C$ is dependent. Since C is minimal dependent, we have $C=\omega(a)\cap C\subset\omega(a)$.

(γ)$\Rightarrow$(β). We may assume that $a\neq 0$ and $b\neq 0$. We put $h(a)=m$ and $h(b)=n$. Since $\omega(a)\cap\omega(b)=\emptyset$ by (γ), there are disjoint independent sets $\{p_1, \dots, p_m\}$, $\{q_1, \dots, q_n\}$ of $\Omega(L)$ with $a=p_1\vee\dots\vee p_m$, $b=q_1\vee\dots\vee q_n$. If $h(a\vee b)<m+n$, then $\{p_1, \dots, p_m, q_1, \dots, q_n\}$ would be dependent, and hence there is $C\in\mathscr{C}(L)$ such that

$$C\subset \{p_1, \dots, p_m, q_1, \dots, q_n\}\subset\omega(a)\cup\omega(b).$$

But, $C\not\subset\{p_1, \dots, p_m\}$ since $\{p_1, \dots, p_m\}$ is independent. Hence, C contains some q_i, so that $C\not\subset\omega(a)$. Similarly, $C\not\subset\omega(b)$, a contradiction. Therefore, $h(a\vee b)= =m+n$.

Lemma 4. *Let* $\omega\subset\Omega(L)$ *and* $p\in\Omega(L)-\omega$. $p\in\mathrm{Cl}(\omega)$ *if and only if there exsits* $C\in\mathscr{C}(L)$ *such that* $p\in C\subset\omega\cup\{p\}$.

Proof. If $p\in\mathrm{Cl}(\omega)$, then there exist $q_1, \dots, q_n\in\omega$ such that $p\leqq q_1\vee\dots\vee q_n$. Let ω_0 be a minimal subset of $\{q_1, \dots, q_n\}$ such that $p\leqq\bigvee(q;\, q\in\omega_0)$. Then, ω_0 must be independent by the minimality. Moreover, for any $q_i\in\omega_0$, $(\omega_0-\{q_i\})\cup\{p\}$ is independent since $p\not\leqq\bigvee(q;\, q\in\omega_0-\{q_i\})$. Therefore, $C=\omega_0\cup\{p\}$ is a circuit and $p\in C\subset\omega\cup\{p\}$.

Conversely, if $C\in\mathscr{C}(L)$ and $p\in C\subset\omega\cup\{p\}$, then we have $p\leqq\bigvee(q;\, q\in C-\{p\})$, and then $p\in\mathrm{Cl}(\omega)$, since $C-\{p\}\subset\omega$.

4. Modularity of *FAC*-lattices. Let L be an atomistic lattice. For $n=1, 2, \dots$, we put

$$\Omega^n = \{p_1\vee\dots\vee p_n;\ p_i\in\Omega(L)\}.$$

Evidently, $\Omega^1=\Omega(L)$, $\Omega^n\subset\Omega^{n+1}$ for every n, and $\bigcup_{n=1}^{\infty}\Omega^n=F(L)-\{0\}$.

For two subsets A, B of L, we write $(A, B)M$ (resp. $(A, B)M^*$) if (a, b) is modular (resp. dual-modular) for all $a\in A$, $b\in B$. The following equivalences are proved in [4] (or [7]).

(1) $(A, \Omega^n)M \Leftrightarrow (A, \Omega^{n-1})M^* \quad (n = 2, 3, \ldots), \quad (A, L)M \Leftrightarrow (A, L)M^*.$

(2) $(\Omega^n, \Omega)M^* \Leftrightarrow (\Omega^{n-1}, \Omega^2)M^* \Leftrightarrow \ldots \Leftrightarrow (\Omega^2, \Omega^{n-1})M^* \quad (n = 3, 4, \ldots).$

(3)
$$(F(L), \Omega^n)M^* \Leftrightarrow (F(L), F(L))M^* \quad (n = 1, 2, \ldots),$$
$$(\Omega^n, F(L))M^* \Leftrightarrow (F(L), F(L))M^* \quad (n = 2, 3, \ldots).$$

If L is an FAC-lattice, then $F(L)=L$, and $(\Omega, L)M^*$ holds by the covering property. For L, we have the following implications:

$(*)$ $\quad (L, L)M^* \Rightarrow \ldots \Rightarrow (\Omega^{m+1}, \Omega)M^* \Rightarrow (\Omega^m, \Omega)M^* \Rightarrow \ldots \Rightarrow (\Omega^2, \Omega)M^*.$

We remark that each of $(\Omega^m, \Omega^n)M^*$, $(\Omega^m, L)M^*$ and $(L, \Omega^n)M^*$ $(m \geqq 2$ and $n \geqq 1)$ is equivalent to some member of $(*)$ by (2) and (3).

Lemma 5. *Let L be an FAC-lattice and let $p \in \Omega(L)$ and $m \geqq 2$. The following statements are equivalent.*

(α) *$(a, p)M^*$ for every $a \in \Omega^m$.*

(β) *If $p \in C \in \mathscr{C}(L)$ and $|C| \leqq m+2$, then for any $q \in C - \{p\}$ there exists $r \in \mathrm{Cl}(C - \{p, q\})$ such that $\{p, q, r\} \in \mathscr{C}(L)$.*

Proof. It follows from [3], Lemma 2 that (α) is equivalent to

(α') $(a, p)P$ for every $a \in \Omega^m$.

($(a, p)P$ means that if $q \in \Omega(L)$ and $q \leqq a \vee p$ then there exists $r \in \Omega(L)$ such that $q \leqq r \vee p$ and $r \leqq a$.) We shall prove (α')$\Rightarrow$(β). Let $p \in C \in \mathscr{C}(L)$, $|C| \leqq m+2$ and $q \in C - \{p\}$. We put $a = \bigvee(r;\ r \in C - \{p, q\})$. Then, $a \in \Omega^m$. We have $p \wedge a = 0$, since $C - \{q\}$ is independent. Similarly, $q \wedge a = 0$. Since $q \in C \subset \omega(a \vee p) \cup \{q\}$, it follows from Lemma 4 that $q \in \mathrm{Cl}(\omega(a \vee p)) = \omega(a \vee p)$, so that $q \leqq a \vee p$. By (α') there exists $r \in \Omega(L)$ such that $q \leqq r \vee p$ and $r \leqq a$. We have $r \neq p, q$ since $p \wedge a = q \wedge a = 0$. Hence, $\{p, q, r\}$ is a circuit and $r \in \omega(a) = \mathrm{Cl}(C - \{p, q\})$.

(β)$\Rightarrow$(α'). Let $a \in \Omega^m$. There is an independent set $\{r_1, \ldots, r_n\}$ with $a = r_1 \vee \ldots \vee r_n$, and then $n \leqq m$. Let $q \leqq a \vee p$ $(q \in \Omega(L))$ and we shall show the existence of $r \in \Omega(L)$ with $q \leqq r \vee p$, $r \leqq a$. We may assume $q \nleqq a$ and $q \neq p$. The set $\{p, q, r_1, \ldots, r_n\}$ is dependent since $q \leqq p \vee r_1 \vee \ldots \vee r_n$. Hence, there is a circuit C such that $\{p, q\} \subset C \subset \{p, q, r_1, \ldots, r_n\}$. By ($\beta$) there exists $r \in \mathrm{Cl}(C - \{p, q\})$ such that $\{p, q, r\} \in \mathscr{C}(L)$. Then, $q \leqq p \vee r$, and we have $r \leqq a$ since $C - \{p, q\} \subset \omega(a)$.

Definition. Let L be a FAC-lattice. A circuit $C \in \mathscr{C}(L)$ is called a *P-circuit* if for every $p, q \in C$ $(p \neq q)$ there exists $r \in \mathrm{Cl}(C - \{p, q\})$ such that $\{p, q, r\} \in \mathscr{C}(L)$. Evidently, if $|C| = 3$ then C is a P-circuit.

Theorem 6. *Let L be an FAC-lattice, and let $m \geqq 2$.*

(i) *L satisfies $(\Omega^m, \Omega)M^*$ if and only if every $C \in \mathscr{C}(L)$ with $|C| \leqq m+2$ is a P-circuit.*

(ii) *L is modular (i.e., $(L, L)M^*$) if and only if every $C \in \mathscr{C}(L)$ is a P-circuit.*

Proof. (i) directly follows from Lemma 5. Since $(L, \Omega)M^* \Leftrightarrow (L, L)M^*$, (ii) follows from (i).

Definition. Let L be a lattice and let $a, b \in L$. (a, b) is called a *distributive pair* (a join-distributive pair in [5]), denoted by $(a, b)D$, when

$$(a \vee b) \wedge x = (a \wedge x) \vee (b \wedge x) \quad \text{for every} \quad x \in L.$$

If L is atomistic, it is easy to verify that $(a, b)D$ is equivalent to the following condition:

If $p \in \Omega(L)$ and $p \leqq a \vee b$ then $p \leqq a$ or $p \leqq b$. Hence, $(a, b)D \Leftrightarrow \omega(a \vee b) = \omega(a) \cup \omega(b)$.

We shall now be interested in an FAC-lattice L satisfying the following condition:

(D) If $C \in \mathscr{C}(L)$ and $|C| \geqq 4$ then for any $p \in C$ there exists $q \in C - \{p\}$ such that $(p, q)D$.

Lemma 7. *If an FAC-lattice L satisfies* (D), *then* $C \in \mathscr{C}(L)$ *is a P-circuit only when* $|C| = 3$.

Proof. Let $C \in \mathscr{C}(L)$ with $|C| \geqq 4$, and let $p \in C$. By (D) there is $q \in C - \{p\}$ such that $(p, q)D$. Then, $\{p, q, r\}$ is not a circuit for any $r \in C - \{p, q\}$, because $r \leqq p \vee q$ implies $r = p$ or $r = q$. Hence, C is not a P-circuit.

Example 8. Let G be a non-oriented finite simple graph, and let $E(G)$ be the set of all edges of G. The cycle matroid $M(G)$ is defined by the collection of independent subsets of $E(G)$, where S is an independent subset if and only if S does not contain a cycle of G ([9], 1.3). A subset C of $E(G)$ is minimal dependent if and only if C is a cycle, and the closure operator in $M(G)$ is defined as follows:

$$x \in \mathrm{Cl}(S) \Leftrightarrow x \in S \quad \text{or there exists a cycle } C \text{ such that} \quad x \in C \subset S \cup \{x\}.$$

It is easy to verify in the same way as in Proposition 1 (ii) that the set

$$L(G) = \{S \subset E(G);\ \mathrm{Cl}(S) = S\}$$

forms an AC-lattice (cf. [9], 3.3). Since $E(G)$ is a finite set, $L(G)$ is an FAC-lattice, and we call it a *bond lattice* associated with G ([6], [8]). We remark that the set $\mathscr{C}(L(G))$ is just the set of all cycles of G.

We shall show that

(G) $(S_1, S_2)D$ in $L(G)$ if S_1 and S_2 has no common vertex.

Let $x \in S_1 \vee S_2 = \mathrm{Cl}(S_1 \cup S_2)$. If $x \notin S_1 \cup S_2$ then there is a cycle C such that $x \in C \subset S_1 \cup S_2 \cup \{x\}$. Since S_1 and S_2 has no common vertex, we have $C \subset S_i \cup \{x\}$ for some i. Then, $x \in \mathrm{Cl}(S_i) = S_i$. Therefore, $(S_1, S_2)D$ holds.

It is easy to show by (G) that any bond lattice satisfies the condition (D). In fact, if C is a cycle with $C \geqq |4|$ and if $x \in C$, then there exists $y \in C$ such that x and y have no common vertex.

Theorem 9. *Let L be an FAC-lattice satisfying* (D) *(for instance, a bond lattice), and let $m \geqq 2$.*

(i) *L satisfies $(\Omega^m, \Omega)M^*$ if and only if there is no $C \in \mathscr{C}(L)$ such that $4 \leqq |C| \leqq m+2$.*

(ii) *L satisfies $(\Omega^m, \Omega)M^*$ but does not satisfy $(\Omega^{m+1}, \Omega)M^*$ if and only if there is $C_0 \in \mathscr{C}(L)$ with $|C_0| = m+3$ and there is no $C \in \mathscr{C}(L)$ such that $4 \leqq |C| \leqq m+2$.*

(iii) *L does not satisfy $(\Omega^2, \Omega)M^*$ if and only if there is $C_0 \in \mathscr{C}(L)$ with $|C_0| = 4$.*

(iv) *L is modular if and only if there is no $C \in \mathscr{C}(L)$ such that $|C| \geqq 4$.*

Proof. Evidently, the statements (i) and (iv) follow from Theorem 6 and Lemma 7, and (ii) and (iii) follow from (i). (The statement (iv) was proved in [8], in case $L = L(G)$.)

We remark that if a graph G is a cycle with $n+3$ edges then $L(G)$ is isomorphic to the lattice given in [4], Example 3.

5. Strongly planar lattices. An AC-lattice L is called *strongly planar* when it satisfies the following condition ([2], (14.3)):

(SP) If $p, q, r \in \Omega(L)$, $a \in L$ and if $p \leqq q \vee a$ and $r \leqq a$ then there exists $s \in \Omega(L)$ such that $p \leqq q \vee r \vee s$ and $s \leqq a$.

It follows from [2], (14.4) that an AC-lattice is strongly planar if either L is modular or the length of L is 3 (i.e., L has 1 and $h(1) = 3$). We call such a lattice a trivial strongly planar lattice. It is well-known that non-modular affine matroid lattices are non-trivial strongly planar lattices (see [2], (18.3) and (14.5)). Here we shall show that if a bond lattice $L(G)$ is strongly planar then it is a trivial one. (Hence, the set of non-modular affine matroid lattices and the set of bond lattices have no common element.)

Firstly we remark that any bond lattice $L = L(G)$ satisfies the following two conditions by the property (G):

(D′) If $C \in \mathscr{C}(L)$ and $|C| \geqq 5$ then there exist three different elements $p, q, r \in C$ such that $(p, q \vee r)D$.

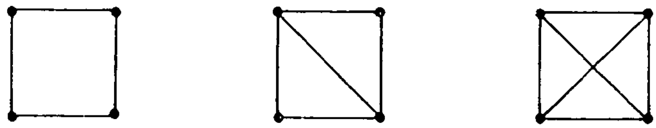
(D″) If $p \in \Omega(L)$, $C \in \mathscr{C}(L)$, $|C| = 4$ and $p \nleqq \vee(q;\ q \in C)$ then there exist $q_1, q_2 \in C$ $(q_1 \neq q_2)$ such that $(p \vee q_1, q_2)D$.

Theorem 10. *Let L be a FAC-lattice satisfying* (D′) *and* (D″) *(for instance, a bond lattice). If L is strongly planar and non-modular then the length of L is* 3.

Proof. Since L is non-modular, by Theorem 6 (ii) there is $C \in \mathscr{C}(L)$ which is not a P-circuit. Then, $|C| \geqq 4$. Suppose $|C| \geqq 5$, then by (D′) there exist three different elements $p, q, r \in C$ such that $(p, q \vee r)D$. Put $a = \bigvee(t;\ t \in C - \{p, q\})$. Since $p \leqq q \vee a$ and $r \leqq a$, by (SP) there exists $s \in \Omega(L)$ such that $p \leqq q \vee r \vee s$ and $s \leqq a$. The set $\{p, q, r\}$ is independent, since C is minimal dependent. Hence, $p \not\leqq q \vee r$, and then we have $s \leqq q \vee r \vee p$ by (α') in the proof of Proposition 1. Since $C - \{q\}$ is independent, we have $p \not\leqq a$, so that $s \neq p$. Hence, by $(p, q \vee r)D$ we have $s \leqq q \vee r$. This implies $p \leqq q \vee r$, a contradiction. Therefore, we obtain that $|C| = 4$.

Next we shall show that $\bigvee(t;\ t \in C) = 1$. Suppose there is $p \in \Omega(L)$ such that $p \not\leqq \bigvee(t, t \in C)$. By (D″) there exist $q_1, q_2 \in C$ such that $(p \vee q_1, q_2)D$. We put $C = \{q_1, q_2, q_3, q_4\}$ and $b = p \vee q_3 \vee q_4$. Since $q_2 \leqq q_1 \vee q_3 \vee q_4 \leqq q_1 \vee b$ and $p \leqq b$, by (SP) there exists $s \in \Omega(L)$ such that $q_2 \leqq q_1 \vee p \vee s$ and $s \leqq b$. We have $q_2 \not\leqq p \vee q_1$, since the set $\{p, q_1, q_2\}$ is independent by $p \not\leqq q_1 \vee q_2$. Hence, we have $s \leqq p \vee q_1 \vee q_2$ by (α'), and then either $s = q_2$ or $s \leqq p \vee q_1$ by $(p \vee q_1, q_2)D$. But, $s \leqq p \vee q_1$ implies $q_2 \leqq q_1 \vee p$, a contradiction. Moreover, since $\{q_2, q_3, q_4\}$ is independent and $p \not\leqq q_2 \vee q_3 \vee q_4$, $\{p, q_2, q_3, q_4\}$ is independent and hence $q_2 \not\leqq b$. Thus, $s = q_2$ contradicts that $s \leqq b$. Therefore, we obtain $\bigvee(t;\ t \in C) = 1$, and then $h(1) = |C| - 1 = 3$.

Corollary 11. *If a bond lattice $L(G)$ is strongly planar and non-modular then G is isomorphic to one of the following three graphs:*



Proof. By the proof of the theorem, there is a cycle $C \subset E(G)$ such that $|C| = 4$ and $\mathrm{Cl}(C) = E(G)$.

References

[1] G. Birkhoff, *Lattice Theory,* Colloq. Publ., Amer. Math. Soc. (Providence, 1967).

[2] F. Maeda and S. Maeda, *Theory of Symmetric Lattices,* Springer-Verlag (Berlin, 1970).

[3] S. Maeda, On finite-modular atomistic lattices, *Algebra Universalis,* **12** (1981), 76—80.

[4] S. Maeda, On modularity in atomistic lattices, in: *Contributions to Lattice Theory* (Proc. Conf. Szeged, 1980), Colloq. Math. Soc. János Bolyai, 33, North-Holland (Amsterdam, 1983), pp. 627—636.

[5] S. Maeda, On distributive pairs in lattices, *Acta Math. Hungar.,* **45** (1985), 133—140.

[6] D. Sachs, Graphs, matroids, and geometric lattices, *J. Combin. Theory,* 9 (1970), 192—199.

[7] N. K. Thakare, M. P. Wasadicar and S. Maeda, On modular pairs in semilattices, *Algebra Universalis,* **19** (1984), 255—265.

[8] M. Tsuchiya, On determination of a graph G whose bond lattice $L(G)$ is modular, *Discrete Math.,* **56** (1985), 79—81.

[9] D. J. A. Welsh, *Matroid Theory,* Academic Press (London, 1976).

DEPARTMENT OF MATHEMATICS
EHIME UNIVERSITY
BUNKYO-CHO, MATSUYAMA, 790, JAPAN

Acta Sci. Math., 51 (1987), 145—151

A note on radical and semisimple classes of topological rings

L. MÁRKI*), R. MLITZ and R. WIEGANDT*)

To the memory of András Huhn

In this note we apply the general Kurosh—Amitsur radical theory presented in [5] to categories of topological rings and deduce characterizations for radical and semisimple classes which go, partly, far beyond the point to which Arnautov and Vodinchar developed these aspects of the radical theory of topological rings. Furthermore, we carry over a characterization of semisimple classes of supernilpotent radicals to the topological case.

1. Let TopR denote the category of Hausdorff topological (associative) rings and all continuous homomorphisms, and $\mathscr{C}$ be a *universal class* in TopR, i.e., a subcategory such that:

(i) if $A\in\mathscr{C}^\circ$ ($\mathscr{C}^\circ$ denotes the class of objects of $\mathscr{C}$) and $B\triangleleft A$ (i.e., B is an ideal of A endowed with the subspace topology) with canonical embedding $\varphi\colon B\to A$, then $B\in\mathscr{C}^\circ$ and $\varphi\in\mathscr{C}$;

(ii) if $A\in\mathscr{C}^\circ$, $\psi\colon A\to C$ is a surjective morphism in TopR, and the topology of C agrees with the quotient topology corresponding to ψ, then $C\in\mathscr{C}^\circ$ and $\psi\in\mathscr{C}$.

In addition, we assume that every morphism in $\mathscr{C}$ admits a unique factorization into the composition of a surjective morphism and a morphism which is a subspace embedding; in other words, if we denote by $\mathscr{E}$ the class of all surjective morphisms and by $\mathscr{M}$ the class of all subspace embeddings in $\mathscr{C}$, then $\mathscr{C}$ admits a unique $(\mathscr{E}, \mathscr{M})$ factorization. Whenever we shall speak of a factorobject $\psi\colon A\to C$ or a subobject $\varphi\colon B\to A$ of an $A\in\mathscr{C}^\circ$, this means that $\psi\in\mathscr{E}$ or $\varphi\in\mathscr{M}$, respectively. We assume also that for every $A\in\mathscr{C}^\circ$, its factorobjects form a complete lattice and its subobjects an inductive set, the latter in the sense that any ascending chain of subobjects has

Received April 1, 1986.

*) Research partially supported by Hungarian National Foundation for Scientific Research grant no. 1813.

a least upper bound. By a trivial object we mean a one-element ring 0, which is necessarily in $\mathscr{C}$. Then $\mathscr{C}$ satisfies all the axioms imposed on the category in [5].

2. In accordance with the terminology in [5], by a radical we mean a mapping ϱ which assings to every $A \in \mathscr{C}^{\circ}$ a factorobject $\varrho_A \colon A \to \varrho(A)$ such that

($\varrho 1$) for every $\varphi \colon A \to C$ from $\mathscr{E}$, there is a $\varrho(\varphi) \colon \varrho(A) \to \varrho(C)$ in $\mathscr{E}$ which makes the diagram

$$\begin{tikzcd} A \arrow[r, "\varphi"] \arrow[d, "\varrho_A"'] & C \arrow[d, "\varrho_C"] \\ \varrho(A) \arrow[r, "\varrho(\varphi)"] & \varrho(C) \end{tikzcd}$$

commutative;

($\varrho 2$) $\varrho_{\varrho(A)} = 1_{\varrho(A)}$ for all $A \in \mathscr{C}^{\circ}$;

($\varrho 3$) $\varrho_A = 1_A$ if and only if for all $0 \neq B \lhd A$, $\varrho(B) \neq 0$.

This notion of radical is formulated in terms of the factorization in $\mathscr{C}$. However, at least if $\mathscr{C}$ is the category of all topological rings and all continuous homomorphisms, then the notion of radical is in fact independent of the factorization. To show this, notice that this category admits two extremal factorizations: (surjective homomorphisms with quotient topology, continuous monomorphisms) and (continuous epimorphisms, extremal monomorphisms with subspace embedding). By Remark 5 to the definition of M-radicals in [5], it suffices to exhibit that these two factorizations admit the same radicals. In view of the considerations there, any ϱ_A obtained in the factorization with the larger class of epimorphisms is an epimorphism of the stricter sense — that is, ϱ_A is necessarily a surjective homomorphism with the quotient topology on $\varrho(A)$. Furthermore, if we have an M-system in the larger sense, i.e., an ideal I with a topology which is maybe finer than the subspace topology, then firstly, the embedding ι of this ideal is algebraically an extremal monomorphism, and secondly, if we consider its factorization in the other sense, then we obtain an

$$\begin{tikzcd} I \arrow[rr, "\iota"] \arrow[dr] & & A \\ & K \arrow[ur, "\varkappa"'] & \end{tikzcd}$$

object K which is the same ideal I with the subspace topology, hence its embedding $\varkappa$ is a monomorphism of the strictest sense. Thus from a radical ideal in the first factorization we obtain a radical ideal in the second one. By this we have the independence we wanted to show.

For a radical ϱ, we denote the radical and the semisimple classes by $\mathbf{R}_\varrho$ and $\mathbf{S}_\varrho$, respectively, i.e.,

$$\mathbf{R}_\varrho = \{A\in\mathscr{C}^0 \colon\ \varrho(A) = 0\}, \quad \mathbf{S}_\varrho = \{A\in\mathscr{C}^0 \colon\ \varrho_A = 1_A\}.$$

We know that each of the radical, the radical class, and the semisimple class determines the other two.

It follows from (ϱ1) that every radical class $\mathbf{R}_\varrho$ is $\mathscr{E}$-closed, i.e., if $\varphi\colon A\to C\in\mathscr{E}$ and $A\in\mathbf{R}_\varrho$ then $C\in\mathbf{R}_\varrho$.

Proposition 1. *Let A be a ring with two topologies σ and τ, $\sigma\leqq\tau$, in $\mathscr{C}^0$. Then $(A,\tau)\in\mathbf{R}_\varrho$ implies $(A,\sigma)\in\mathbf{R}_\varrho$ and $(A,\sigma)\in\mathbf{S}_\varrho$ implies $(A,\tau)\in\mathbf{S}_\varrho$.*

Proof. The first claim follows from the $\mathscr{E}$-closedness of $\mathbf{R}_\varrho$, and implies by (ϱ3) the second one.

For an arbitrary subclass $\mathbf{R}\subseteqq\mathscr{C}^0$ and any $A\in\mathscr{C}^0$ we put $\mathbf{R}A=\sum(B\lhd A,\ B\in\mathbf{R})$ endowed with the subspace topology. Clearly, $\mathbf{R}A\lhd A$.

In view of [5] Proposition 4.2 and the characterization II. 1—2 of radical classes in ARNAUTOV and VODINCHAR [4], our radicals are the same as those in [4]. (Notice that our definition makes no allusion to the closedness of the largest radical ideal!) Therefore $\mathbf{R}_\varrho A$ is the kernel of ϱ_A (hence a closed ideal of A) and $\mathbf{R}_\varrho A\in\mathbf{R}_\varrho$. Thus every radical in $\mathscr{C}$ is attainable in the sense of [5]. [5] Theorem 3.1 and Proposition 4.1 yield now the following characterization of radicals.

Proposition 2. *A mapping ϱ which assigns to each $A\in\mathscr{C}^0$ a factorobject $(\varrho_A, \varrho(A))$ is a radical if and only if it satisfies conditions (ϱ1), (ϱ2), and*

(ϱ3) for every $A\in\mathscr{C}^0$ there is an ideal $I\lhd A$ such that $\varrho(I)=0$, ϱ_A is the canonical factor $A\to A/\bar{I}$ (where $\bar{I}$ denotes the closure of I) and for all $J\lhd A$, $\varrho(J)=0$ implies $J\subseteqq\bar{I}$.*

Remark. Notice that in condition (ϱ3*) it is not required that the ideal I be closed; however, it can always be chosen to be closed, as was shown above.

3. From [5] we obtain now three characterizations of radical classes. The first of them is a simple transcription of [5] Proposition 3.4, the latter two follow from Remark 3.7, which is easily seen to apply in our case. Therefore we shall give no proof here. Characterization (I) is just the definition of radical classes in ARNAUTOV and VODINCHAR [4]. Notice that the closedness of the largest radical ideal is imposed only in characterization (I).

Theorem 3. *A class $\mathbf{R}\subseteqq\mathscr{C}^0$ is a radical class if and only if it satisfies*

(I) (R2) $\mathbf{R}$ *is $\mathscr{E}$-closed,* (R4) *in every* $A \in \mathscr{C}^\circ$ *there is an ideal* $\mathbf{R}(A)$ *such that* $\mathbf{R}(A) \in \mathbf{R}$ *and* $I \subseteqq \overline{\mathbf{R}(A)}$ *for all* $I \lhd A$, $I \in \mathbf{R}$,

(R5) $\mathbf{R}(A/\overline{\mathbf{R}(A)}) = 0$,

and $\mathbf{R}(A)$ *in* (R4) *is closed in* A; *or*

(II) (R2), (R4), *and*

(R3) *if* $I \lhd A$, $I \in \mathbf{R}$ *and* $A/\bar{I} \in \mathbf{R}$ *then* $A \in \mathbf{R}$; *or*

(III) (R2), (R3), *and*

(R4′) *in every object, the union of any chain of ideals from* $\mathbf{R}$ *belongs to* $\mathbf{R}$.

4. Let S be a subclass of $\mathscr{C}^\circ$ which is closed under subdirect products (the topology on a subdirect product is the subspace topology of the product topology; of course, we consider only those subdirect products which are in $\mathscr{C}^\circ$). Then every $A \in \mathscr{C}^\circ$ has a largest factorobject in S; we denote by $\mathrm{S}(A)$ the kernel ideal belonging to this factor (then $\mathrm{S}(A)$ is necessarily closed).

By Theorem 1 in Arnautov [3] every radical in $\mathscr{C}$ has the A—D—S property, i.e., for any radical class $\mathbf{R}$ and any $A, B \in \mathscr{C}^\circ$, $B \lhd A$, we have $\mathbf{R}B \lhd A$. Consequently, every semisimple class in $\mathscr{C}$ is hereditary (with respect to ideals).

We also have the obvious characterization (see Arnautov and Vodinchar [4]): $\mathrm{S} \subseteqq \mathscr{C}^\circ$ is a semisimple class if and only if, for all $A \in \mathscr{C}^\circ$,

$$A \in \mathrm{S} \Leftrightarrow \forall B \lhd A\colon\ B \neq 0 \Rightarrow B \text{ has a non-zero factor in S.}$$

Theorem 3.6 from [5] translates into the following:

A class $\mathrm{S} \subseteqq \mathscr{C}^\circ$ *is a semisimple class if and only if*

(S3) S *is closed under subdirect products,*

(S4′) S *is regular, i.e., if* $A \in \mathrm{S}$ *and* $0 \neq B \lhd A$ *then* B *has a non-zero factor in* S,

(S6) *if* $\psi\colon A \to B$ *is a surjective continuous homomorphism with* $\operatorname{Ker}\psi \in \mathrm{S}$ *and* $B \in \mathrm{S}$ *then* $A \in \mathrm{S}$,

(S7) $\mathrm{S}(\mathrm{S}(A)) \lhd A$ *for all* $A \in \mathscr{C}^\circ$.

In fact, here (S7) follows from the other conditions, and (S3) can be weakened to the coinductive property

(S3′) if (I_α) is a descending chain of closed ideals in $A \in \mathscr{C}^\circ$ such that $A/I_\alpha \in \mathrm{S}$ for all α, then $A/\bigcap I_\alpha \in \mathrm{S}$.

Theorem 4. *A class* $\mathrm{S} \subseteqq \mathscr{C}^\circ$ *is a semisimple class if and only if it satisfies* (S3′), (S4′) *and* (S6).

Lemma 5. *Suppose that* $\mathrm{S} \subseteqq \mathscr{C}^\circ$ *satisfies* (S3′), (S4′), (S6). *If* $I \lhd A \in \mathrm{S}$ *and* $I^2 = 0$, *then also* $I \in \mathrm{S}$.

Proof. At first we shall prove the validity of the weaker statement: if $I \lhd A \in \mathrm{S}$ and $A^2 = 0$, then $I \in \mathrm{S}$. By condition (S3′) Zorn's lemma is applicable, so there exists a closed ideal J of I which is minimal with respect to the property $I/J\in\mathrm{S}$. If $J\neq 0$ then $J\lhd A$ because $A^2=0$, hence by (S4′) there is a closed ideal $K\lhd J$ such that $0\neq J/K\in\mathrm{S}$. Again we have $K\lhd I$, K is closed in I, and so $(I/K)/(J/K)\cong\cong I/J\in\mathrm{S}$ holds, hence condition (S6) yields $I/K\in\mathrm{S}$. By the minimality of J it follows now $K=J$, i.e., $J/K=0$, a contradiction. Thus $J=0$ and $I\in\mathrm{S}$.

Now we turn to the proof of the general case of Lemma 5, and choose J again as before. If $J\lhd A$ then, as above, we conclude $I\in\mathrm{S}$. Suppose therefore that J is not an ideal of A. Then there exists an element $a\in A$ such that, say, $aJ\nsubseteq J$. Now we have $0\neq(aJ+J)/J\lhd I/J\in\mathrm{S}$ and $(I/J)^2=0$, hence the foregoing consideration yields that $(aJ+J)/J\in\mathrm{S}$. Furthermore, it is easy to check that the mapping

$$\varphi\colon J\to(aJ+J)/J \quad \text{defined by} \quad j\mapsto aj+J$$

is a continuous surjective homomorphism (J and $(aJ+J)/J$ have the subspace topology induced by I and I/J, respectively) and that $\operatorname{Ker}\varphi$ is a closed ideal not only of J but also of I. Also, the algebraic isomorphism $J/\operatorname{Ker}\varphi\to(aJ+J)/J$ is easily seen to be continuous, therefore $J/\operatorname{Ker}\varphi\in\mathrm{S}$ by (S6). Now $(I/\operatorname{Ker}\varphi)/(J/\operatorname{Ker}\varphi)\cong\cong I/J\in\mathrm{S}$, hence by (S6) we conclude that $I/\operatorname{Ker}\varphi\in\mathrm{S}$. Then by the minimality of J we have $J=\operatorname{Ker}\varphi$ and so $aJ+J=J$, a contradiction. Hence $J\lhd A$, and the lemma is proven.

Proof of Theorem 4. We have already seen that the conditions (S3′), (S4′), (S6) are necessary. In view of an observation made at the beginning of section 4, the sufficiency will be proven if we exhibit that the converse of (S4′) holds. So, let $A\in\mathscr{C}^0$ be such that every non-zero ideal of A has a non-zero factor in S. Then by (S3′) and (S4′) there exists a closed ideal $I\lhd A$ such that $A/I\in\mathrm{S}$ and I is minimal with respect to this property. We shall show that $I=0$ and so $A\in\mathrm{S}$. Assume that $I\neq 0$. Applying (S4′) to $I\lhd A$ and (S3′) to I, we obtain a closed ideal $J\lhd I$ such that $0\neq I/J\in\mathrm{S}$ and J is minimal with respect to this property. We claim that $J\lhd A$. Assume that this is not the case and that $aJ\nsubseteq J$ for an element $a\in A$. Then we have, as in Lemma 5, a continuous surjective homomorphism

$$\varphi\colon J\to(aJ+J)/J$$

with $\operatorname{Ker}\varphi\lhd I$, and by Lemma 5 we have $(aJ+J)/J\in\mathrm{S}$. Now we proceed exactly as in the proof of the general case in Lemma 5, and arrive at $J\lhd A$. Then (S6) together with $(A/J)/(I/J)\cong A/I\in\mathrm{S}$ and $I/J\in\mathrm{S}$ yields $A/J\in\mathrm{S}$. By the minimality of I we have now $I=J$, contrary to the assumption $I/J\neq 0$. Thus the case $I\neq 0$ is impossible, and the proof of the theorem is complete.

Remark. If all rings in $\mathscr{C}$ are compact or linearly compact in the narrow sense, then by Ánh [2] we also know that S is a semisimple class if and only if it satisfies (S6), is hereditary, and is closed under inverse limits.

5. We also have the following characterization for pairs of corresponding radical and semisimple classes, which is a transcription of [5] Theorem 3.5.

Theorem 6. *A pair* $(\mathbf{R}, \mathrm{S})$ *of subclasses of* $\mathscr{C}^\circ$ *is a pair of corresponding radical and semisimple classes if and only if*

(α) $\mathbf{R} \cap \mathrm{S} = \{0\}$,

(β'') *if* $A \in \mathbf{R}$ *then* A *has no non-zero factors from* S,

(γ) *if* $A \in \mathrm{S}$ *then* A *has no non-zero ideals from* $\mathbf{R}$,

(δ) *each* $A \in \mathscr{C}^\circ$ *has a closed ideal* I *such that* $I \in \mathbf{R}$ *and* $A/I \in \mathrm{S}$.

6. Finally we are going to characterize semisimple classes of supernilpotent radicals.

Lemma 7. *Let* A *be a topological ring and* $I \lhd A$ *(not necessarily closed). Further, let* K *be an ideal of* A *which is maximal relative to* $I \cap K = 0$. *Then* $I \cap \overline{K}$ *is contained in the annihilator of* I *in* A.

Proof. If K is closed then there is nothing to prove. If K is not closed then let $a \in I \cap \overline{K}$ be any non-zero element. Now every neighbourhood U_a of a such that $0 \notin U_a$ contains an element $x \in K$, so we have

$$Ix + xI \subseteqq IK + KI \subseteqq I \cap K = 0.$$

Hence each neighbourhood of a contains a two-sided annihilator of I. By the continuity of multiplication also a annihilates I. Since a was arbitrary, we are done.

Corollary 8. *Let* S *be a regular class of topological rings which contains no non-trivial zero-rings. If* $I \lhd A$ *and* $I \in \mathrm{S}$, *then any ideal* K *of* A *which is maximal relative to* $I \cap K = 0$, *is closed in* A.

Proof. If K is not closed then by Lemma 7 $0 \neq I \cap \overline{K} \subseteqq \mathrm{ann}_A I$. Hence $I \cap K$ is a zero-ring, and at the same time an ideal of I. Since S is regular, a non-trivial homomorphic image of $I \cap \overline{K}$ is in S, a contradiction.

Recall that a radical class is said to be *supernilpotent* if it is hereditary and contains all nilpotent rings, and that a class $\mathbf{C}$ of (topological) rings is said to be *closed under essential extensions* if $A \in \mathbf{C}$ whenever $\mathbf{C}$ contains an essential ideal of A.

Theorem 9. *A class* $\mathrm{S} \subseteqq \mathscr{C}^\circ$ *is the semisimple class of a supernilpotent radical if and only if* S *is regular, closed under subdirect products and essential extensions, and consists of semiprime rings.*

Proof. The standard proof for the discrete case (see Anderson and Wiegandt [1]) works, as the ideal K in the Corollary is closed.

In [4] ARNAUTOV and VODINCHAR proved the following strong result: in the universal class of all (Hausdorff) topological rings a hereditary radical class is either supernilpotent or subidempotent (that is, it consists of idempotent rings). This gives rise to the following

Problem. Characterize the semisimple classes of hereditary radicals of topological rings (by characterizing the semisimple classes of subidempotent radicals and using the above quoted result of Arnautov and Vodinchar).

Acknowledgement. The authors thank Dr's K. Beidar, S. Glavatsky and A. Mikhalev for correcting an error in the formulation of condition (S6).

References

[1] T. ANDERSON and R. WIEGANDT, On essentially closed classes of rings, *Ann. Univ. Sci. Budapest.*, **24** (1981), 107—111.

[2] P. N. ÁNH, On semisimple classes of topological rings, *Ann. Univ. Sci. Budapest.*, **20** (1977), 59—70.

[3] V. I. ARNAUTOV, Complementary radicals in topological rings, *Mat. Issled.*, 3 (1968), no. 2, 16—30. (Russian)

[4] V. I. ARNAUTOV and M. I. VODINCHAR, Radicals of topological rings, *Mat. Issled.*, 3 (1968), no. 2, 31—61. (Russian)

[5] L. MÁRKI, R. MLITZ and R. WIEGANDT, A general Kurosh—Amitsur radical theory, preprint

(L. M., R. W.)
MATHEMATICAL INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES
PF. 127
1364 BUDAPEST, HUNGARY

(R. M.)
INSTITUT FÜR ANGEWANDTE UND
NUMERISCHE MATHEMATIK, TU WIEN
WIEDNER HAUPTSTR. 8—10
1040 WIEN, AUSTRIA

Acta Sci. Math., 51 (1987), 153—162

Distributive congruence lattices of finite algebras

P. P. PÁLFY

To the memory of András Huhn

The most famous open problem in universal algebra is the representation of finite lattices as congruence lattices of finite algebras. The general question is very hard, and in essence it is a group theoretic problem (see [12], [10], [9]). Though, representing finite *distributive* lattices is an easy job. Perhaps the most standard way to do this is by starting with a boolean lattice B containing the given finite distributive lattice D and then adding the closure operation $f\colon B\to B$, defined by $f(x)=\bigwedge\{y\in D\colon y\geqq x\}$; it is easy to see that $\operatorname{Con}(B;\vee,\wedge,f)\cong D$. Another result which shows that it is extremely easy to find congruence representations for finite distributive lattices, due to QUACKENBUSH and WOLK [14], states that for any finite distributive sublattice D of $\operatorname{Eq}(A)$ — the lattice of equivalence relations over the set A — containing the equality and the total relation, some (unary) operations can be defined on A so that the congruences will be exactly the members of D. It was shown by P. PUDLÁK [13] that only the distributive finite lattices have this property, i.e. for any other finite lattice there is a representation by equivalences which is not the congruence lattice of any algebra defined on the given set.

In this paper we deal with the problem of representing *all* finite distributive lattices as congruence lattices of finite algebras belonging to some given class of algebras. For completeness we will cite some known results as well. The answer is positive for lattices (DILWORTH), groups (SILCOCK [18]), solvable groups (Theorem 2.2), modules (trivial, see Proposition 4.1), 2-unary algebras (Theorem 5.3), transitive permutation groups regarded as unary algebras (TŮMA [19], see also Proposition 5.5), algebras of any given type except the 1-unary (Corollary 6.1). There exist finite distributive lattices which are not representable as the congruence lattice of a finite ring (Proposition 3.1), of a 1-unary algebra (Corollary 5.2).

Received April 24, 1986.

Research supported by Hungarian National Foundation for Scientific Research grant No. 1813.

Throughout the paper D will stand for a finite distributive lattice with minimal element 0, J will denote the set of join-irreducible elements of D (0 is not regarded to belong to J), and $n=|J|$, the length of D.

1. Lattices. For lattices we only recall some well-known results. The most basic one is the following, which was first obtained by R. P. Dilworth (mentioned in [1] without proof), see also G. GRÄTZER and E. T. SCHMIDT [4].

1.1. Theorem. *Every finite distributive lattice is isomorphic to the congruence lattice of a finite lattice.*

The congruence lattice of a finite modular lattice is always boolean, but this is no longer true for *infinite* modular lattices. The following remarkable result is worth digressing from our topic of distributive congruence lattices of *finite* algebras.

1.2. Theorem (E. T. SCHMIDT [16]). *Every finite distributive lattice is isomorphic to the congruence lattice of a modular lattice.*

R. FREESE [3] proved that the lattices can be chosen to be finitely generated.

For more detailed discussion we refer the reader to E. T. SCHMIDT's lecture notes [17].

2. Groups. The question for groups was first dealt with by J. KUNTZMANN [8] in 1947, but his construction was not correct (see [15], p. 101). The solution came thirty years later:

2.1. Theorem (H. L. SILCOCK [18]). *Every finite distributive lattice is isomorphic to the congruence lattice (i.e. the lattice of normal subgroups) of a finite group.*

Silcock's construction is based on wreath products of nonabelian simple groups, but he also announced the solvable version of the result, see [18], p. 371. However, his construction of solvable groups with given distributive lattice of normal subgroups is rather complicated and has not been published. Since we deem our construction quite natural, we prove it here:

2.2. Theorem. *Every finite distributive lattice is isomorphic to the lattice of normal subgroups of a finite solvable group.*

Proof. All groups which will appear in the construction will have the property that in any chief series (i.e. maximal chain of normal subgroups) $1=N_0<N_1<\dots$ $\dots<N_{n-1}<N_n=G$ the chief factors N_{i+1}/N_i $(i=0, 1, \dots, n-1)$ are elementary abelian p_i-groups for pairwise different prime numbers p_i (p_i will be called the *characteristic* of N_{i+1}/N_i). Then any chief series of G has this property by the Jordan—Hölder theorem. It implies that in any factor group G/N, no two minimal normal subgroups M_1/N and M_2/N can be isomorphic, as $N<M_1<M_1M_2$ is extendable

to a chief series of G. This property will enable us to make use of a theorem of R. KOCHENDÖRFFER [7], which ensures the existence of a faithful irreducible representation of G/N over the p-element field for any prime p not dividing the order of G/N. (This is in fact a strong sufficient condition derived from the necessary and sufficient condition given by Kochendörffer.) In other words this means that there exists an elementary abelian p-group A and a homomorphism $\varphi\colon G\to\mathrm{Aut}(A)$ with $\mathrm{Ker}\,\varphi=N$ such that there are no nontrivial subgroups of A invariant for the group of automorphisms $\varphi(G)$.

The construction will go by induction on the length of the finite distributive lattice D. Let a be an atom in D, $b=\bigvee\{x\in D\colon x\wedge a=0\}$, then $a\wedge b=0$. Let D_1 be the distributive lattice $\{x\in D\colon x\geqq a\}$. By the induction hypothesis, there exists a finite solvable group G with chief factors of different characteristics whose lattice of normal subgroups is isomorphic to D_1. Let B be the normal subgroup of G corresponding to $a\vee b\in D_1$. Choose a prime p not dividing the order of G. By the cited result of Kochendörffer, there exists a faithful irreducible representation of G/B over the p-element field, i.e. we have an elementary abelian p-group A and a homomorphism $\varphi\colon G\to\mathrm{Aut}(A)$ with $\mathrm{Ker}\,\varphi=B$ and $\varphi(G)$ acting irreducibly on A. Now form the semidirect product $\bar{G}=AG$ with respect to φ. Then the irreducibility of the representation φ implies that A is a minimal normal subgroup of $\bar{G}$, and by the choice of p, the characteristics of the factors in a chief series of $\bar{G}$ are also pairwise different.

Now let $N\lhd\bar{G}$. Since A is a minimal normal subgroup of $\bar{G}$, it follows that either $N\geqq A$ or $N\cap A=1$. In the first case, $N=A(N\cap G)$ with $N\cap G\lhd G$. In the second, $N\leqq C_{\bar{G}}(A)=A\times B$, and as N contains no elements of orders divisible by p, we have $N\leqq B$. Conversely, if $N_1\lhd G$, then $AN_1\lhd\bar{G}$; if $N_1\lhd G$ and $N_1\leqq B$ then $N_1\lhd\bar{G}$. Hence the lattice of normal subgroups of $\bar{G}$ is isomorphic to D.

The proof was based on an idea from the author's earlier work [11].

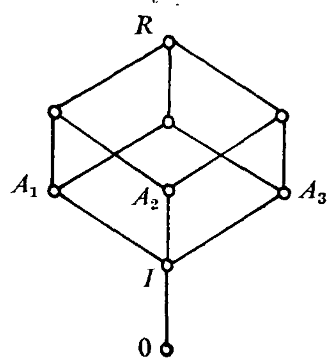
Solvability in Theorem 2.2 cannot be replaced by nilpotency:

2.3. Proposition. *If the lattice of normal subgroups of a finite nilpotent group G is distributive, then G is cyclic, and the lattice is a direct product of chains.*

Proof. Let $\Phi(G)$ denote the Frattini subgroup of G. Then $G/\Phi(G)$ is abelian, it is a direct product of some cyclic groups of prime orders. If two generators had the same prime order then the lattice of (normal) subgroups of $G/\Phi(G)$ would not be distributive. Hence $G/\Phi(G)$ is cyclic, therefore G is a cyclic group itself. Now the lattice of normal subgroups of G is isomorphic to the lattice of divisors of the order of G.

3. Rings. For rings the answer to our representation problem is negative.

3.1. Proposition. *No finite associative ring has congruence lattice (i.e. lattice of ideals) isomorphic to*



Before proving this proposition, let us note that the finiteness of the ring is a crucial requirement, since the following is true:

3.2. Theorem (K. H. Kim and F. W. Roush [6]). *Every finite distributive lattice is isomorphic to the lattice of ideals of some (regular) ring.*

Proof of Proposition 3.1. By way of contradiction assume that a finite ring R has the indicated lattice of ideals. Let I, A_1, A_2, A_3 be ideals of R as shown. We shall reach the final contradiction in several steps. The first observation is obvious:

1) $R/I=A_1/I\oplus A_2/I\oplus A_3/I$, the direct summands are simple rings, in particular any of them is either a ring with unit or a zeroring of prime order.

2) $I^2=0$.

Let $J(R)$ be the Jacobson radical of R. Since $R/J(R)$ is semisimple, its ideal lattice is boolean. Hence we have $J(R)\supseteqq I$. So I is a nilpotent ideal, and by the minimality of I it follows that $I^2=0$.

3) If A_i/I is a ring with unit then there is an idempotent $e_i\in A_i$ for which e_i+I is the unit of A_i/I.

Let $a+I$ be the unit of A/I (for simplicity we leave out the index i in the proofs of steps 3, 4 and 5). Then $a^2=a+t$ for some $t\in I$. Now the required element is $e=a+t-2at$.

4) In the situation of step 3, either $e_iI=0$ or $e_ix=x$ for all $x\in I$.

We show that eI is an ideal of R. Obviously, $eIR\subseteqq eI$. On the other hand, since e is central in R/I, $ReI\subseteqq(eR+I)I\subseteqq eI+0=eI$. If $eI\neq 0$ then $eI=I$ and by the finiteness of I the left multiplication by e induces a permutation on I. Since e is idempotent, it is the identical permutation.

5) In the situation of step 3, $e_iI=0$ and $Ie_i=0$ cannot hold simultaneously. Let $x\in A$ be an arbitrary element. Then $x-ex\in I$ and $Ie=0$ implies $(x-ex)e=0$. So $xe=exe$ and symmetrically $ex=exe$, therefore $xe=ex$ for all $x\in A$. Hence $eA=Ae$ is both right and left ideal of R. We have $eA\neq 0$, since $e=e^2\in eA$. However, $eA\cap I=0$, as for $ex\in eA\cap I$ it follows that $ex=e(ex)\in eI=0$. This contradiction proves that $eI=0=Ie$ is impossible.

6) If A_i/I and A_j/I are rings with unit $(i\neq j)$ then $e_iI=I$ and $e_jI=I$ cannot hold simultaneously.
Otherwise, by step 4 we would have $e_ix=x=e_jx$ for all $x\in I$, hence $e_i-e_j\in \mathrm{Ann}_l I$. This left annihilator is an ideal, but the least ideal of R containing e_i-e_j is A_i+A_j which contains e_i and e_j as well, a contradiction.

7) If A_i/I is a ring with unit and A_j/I is a zeroring then $e_iI=0$.
If not, then $e_ix=x$ for all $x\in I$ by step 4. In particular, $e_iI=e_iA_j=I$. Since A_j/I is a zeroring, $A_j^2\subseteq I$, $IA_j^2\subseteq I^2=0$ hence also $IA_j=0$. Define $B=\{y\in A_j\colon e_iy=0\}$. B is an ideal of R, since $RB\subseteq RA_j\subseteq A_j$, $e_iRB\subseteq (Re_i+I)B\subseteq Re_iB+IA_j=0$ and $BR\subseteq A_jR\subseteq A_j$, $e_iBR=0$. For the left multiplication by e_i, $\lambda\colon A_j\to A_j$, $\lambda(y)=e_iy$ we have $\lambda^2=\lambda$, hence $A_j=\mathrm{Ker}\,\lambda\oplus \mathrm{Im}\,\lambda=B\oplus I$. This is a contradiction.

8) A_i/I and A_j/I $(i\neq j)$ cannot be both zerorings.
Since R is directly indecomposable, its additive group is a p-group for some prime p. Hence A_i/I and A_j/I would be isomorphic zerorings and thus there would be another $p-1$ ideals between I and A_i+A_j.

9) Conclusion. We have already eliminated all possible cases. If all of A_1/I, A_2/I, A_3/I are rings with unit then step 6 implies that $e_iI=0$ for at least two indices i and symmetrically $Ie_j=0$ for at least two j's. Hence for some i we have $e_iI=0=Ie_i$ contrary to step 5. By step 8, there cannot be more than one zerorings among the direct summands. If A_i/I is a ring with unit and A_j/I is a zeroring, then step 7 gives that $e_iI=0$ and by symmetry $Ie_i=0$ as well, again a contradiction by step 5.

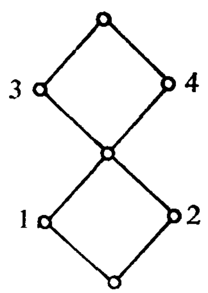
4. Modules. First we present a very elementary construction of a module with given finite distributive lattice of submodules. A similar result for modules over group algebras was obtained by S. M. Vovsi [20].

4.1. Proposition. *Every finite distributive lattice is isomorphic to the congruence lattice (i.e. lattice of submodules) of a finite module.*

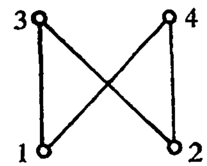
Proof. Recall the definitions of n and J in the introduction. Let F be an arbitrary finite field, $M=F^n$, R the subspace of n by n matrices over F with row and column indices from J spanned by the elementary matrices e_{ij} for $i\leq j$ in J (clearly, R is a ring), and let R act on M in the obvious way. We claim that the submodules of the R-module M are the F-subspaces spanned by the vectors e_i, $i\in I$, for a hereditary subset I of J. Indeed, let N be a submodule, $m\in N$, $m=\sum_{i\in J} f_ie_i$, $f_i\in F$, then for $f_i \neq 0$ we get $e_i = (1/f_i) e_{ii} m \in N$. Hence N is spanned by some of the e_i's. Since $e_{ij} \in R$ for $i \leqq j$ (in J), it follows that $I = \{i \in J \colon e_i \in N\}$ is hereditary. The converse is obvious.

4.2. Example.

D



J



R is the ring of all matrices of the form

$$\begin{bmatrix} a_{11} & 0 & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & 0 \\ 0 & 0 & 0 & a_{44} \end{bmatrix} \quad (a_{ij} \in F)$$

Notice that the ring R depends on the lattice D. If we would like to have modules over the same ring we could take the direct sum of all these rings, or the free (non-commutative) ring with infinitely many generators. It would be desirable to choose a finite ring, however; it is not possible.

4.3. Proposition. *For any finite ring R, there exists a finite distributive lattice D (in fact a chain) such that no finite R-module has lattice of submodules isomorphic to D.*

Proof. Let $J(R)$ be the Jacobson radical of R. Since R is finite, $J(R)$ is nilpotent, i.e. $J(R)^r = 0$ for some positive integer r, and $R/J(R)$ is semisimple. Suppose that the submodules of the R-module M form a chain $0 = M_0 < M_1 < \ldots < M_{n-1} < < M_n = M$. On one hand, $J(R)$ annihilates M_{i+1}/M_i $(i = 0, 1, \ldots, n-1)$, since it is a minimal R-module; on the other hand, if M_j/M_i $(0 \leqq i < j \leqq n)$ is annihilated by $J(R)$ then it can be regarded as an $R/J(R)$-module, hence it is semisimple, which forces $j = i+1$. Thus we have $J(R) \cdot M_{i+1} = M_i$ $(i = 0, 1, \ldots, n-1)$. Now it follows by induction that

$$J(R)^r \cdot M = \begin{cases} M_{n-r} & \text{if } \ n-r > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Since $J(R)^r = 0$, we have $n - r \leqq 0$, $n \leqq r$. Hence no chain longer than r is representable as the congruence lattice (i.e. lattice of submodules) of an R-module.

5. Unary algebras. The 1-unary algebras with distributive congruence lattices have been determined by D. P. Egorova [2]. In order to formulate her result we need some notation. Let $(A; f)$ be a 1-unary algebra, for $a \in A$ we put $f^0(a) = a$,

$f^1(a)=f(a)$, $f^{i+1}(a)=f(f^i(a))$, $i=1,2,\ldots$. On the set $\mathbf{Z}\cup\{\infty\}$ we define the operation f by $f(n)=n+1$ for $n\in\mathbf{Z}$ and $f(\infty)=\infty$. According to [2] the isomorphism types of 1-unary algebras with distributive congruence lattices are the following:

(1) $$\langle a \mid f^{t+r}(a) = f^t(a)\rangle, \quad t \geqq 0, r \geqq 1;$$

(2) $$\langle a, b \mid f^{t+r}(a) = f^t(a), f^s(b) = b\rangle,$$
$$t \geqq 0, \quad r \geqq 1, \quad s \geqq 1 \quad \text{and} \quad \text{g.c.d.}(r, s) = 1;$$

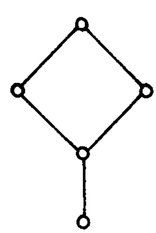
(3) four infinite algebras: $\mathbf{Z}\cup\{\infty\}$, $\mathbf{Z}$, $\mathbf{N}\cup\{\infty\}$, $\mathbf{N}$.

It is quite easy to determine their congruence lattices. Let $C(t)$ denote the chain of length t (i.e. having $t+1$ elements), $D(r)$ the lattice of divisors of r, and $L+1$ the lattice obtained from the lattice L by adding a new maximal element to it. Restricting our attention to finite algebras, we obtain that the congruence lattice in case (1) is isomorphic to $C(t)\times D(r)$, and in case (2) to $C(t)\times(D(rs)+1)$. Hence we have:

5.1. Proposition. *If the finite distributive lattice D is isomorphic to the congruence lattice of a finite 1-unary algebra, then either D is a direct product of chains or $D\cong C_0\times(C_1\times\ldots\times C_k+1)$ for some finite chains $C_0, C_1, \ldots, C_k$.*

Now it is easy to exhibit a finite distributive lattice which is not representable as the congruence lattice of a (finite) 1-unary algebra, cf. [5], p. 209, where this example is credited to J. Johnson and R. L. Seifert.

5.2. Corollary. *No finite 1-unary algebra has congruence lattice isomorphic to*



On the other hand, two unary operations already suffice.

5.3. Theorem. *Every finite distributive lattice is isomorphic to the congruence lattice of a finite 2-unary algebra.*

Proof. For the sake of simplicity suppose $J=\{1, 2, \ldots, n\}$, and let $J'=\{0, 1, \ldots, n\}$. Choose pairwise different primes $p_1, p_2, \ldots, p_n>n$, and let $p_0=1$. For the base set of the algebra we take $A=\{(j, k)\colon j\in J',\ 0\leqq k\leqq p_j-1\}$, and we define the two unary operations by

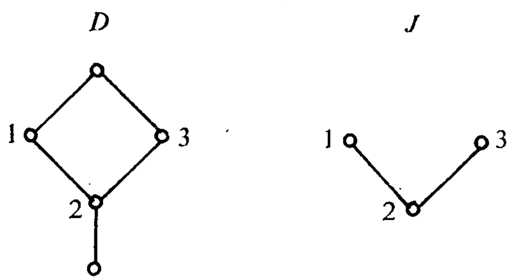
$$f((j,k)) = \begin{cases} (j, k+1) & \text{if } \; k \leqq p_j - 2 \\ (j, 0) & \text{if } \; k = p_j - 1 \end{cases}$$

and

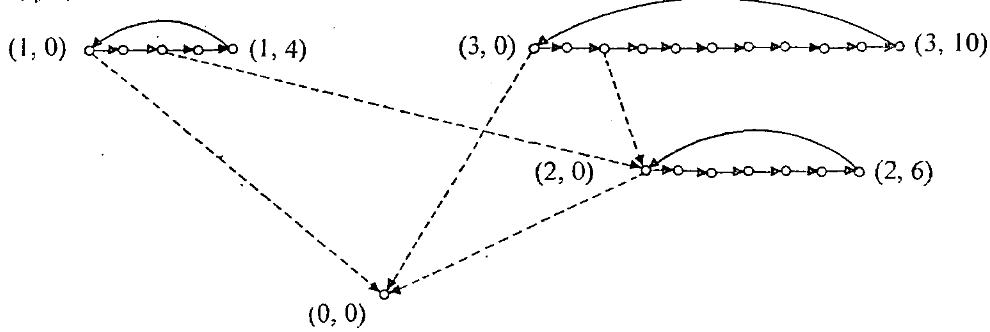
$$g((j,k)) = \begin{cases} (k, 0) & \text{if } \; k = 0 \; \text{ or } \; k < j \; \text{ in } \; J \\ (j, k) & \text{otherwise.} \end{cases}$$

We claim that any nontrivial congruence of the algebra $(A; f, g)$ has one nontrivial class $\{(j,k)\colon j \in I',\ 0 \leqq k \leqq p_j - 1\}$ where $I' = \{0\} \cup I$ for some hereditary subset I of J, the other classes are singletons. The proof of this statement is straightforward and left to the reader. Hence we see that $\mathrm{Con}\,(A; f, g) \cong D$.

5.4. Example.



$(A; f, g)$



Notation: $f \longrightarrow$, $g \dashrightarrow$ (fixed points are not denoted)

P. P. Pálfy and P. Pudlák [12] showed that every finite lattice is representable as a congruence lattice of a finite algebra if and only if every finite lattice is isomorphic to an interval in the subgroup lattice of a finite group. (In fact, the interval $[H, G]$ is isomorphic to the congruence lattice of the unary algebra on the set of left cosets of the subgroup H in the group G, with the operations being the permutations defined by left multiplications by the elements of G.) For finite distributive lattices we construct suitable intervals by applying Silcock's theorem (see 2.1). J. Tůma [19] has given another construction recently.

5.5. Proposition. *Every finite distributive lattice is isomorphic to the congruence lattice of a unary algebra where the operations form a transitive permutation group.*

Proof. Let G be a finite group with $\mathrm{Con}\, G \cong D$ (see Theorem 2.1). Take the diagonal subgroup $\Delta = \{(g, g): g \in G\}$ of $G \times G$. It is easy to prove that the subgroups K containing Δ have the form $K = \Delta \cdot (K_1 \times 1)$, where $K_1 \lhd G$, $K_1 \times 1 = K \cap (G \times 1)$. Hence the interval $[\Delta, G \times G] \cong \mathrm{Con}\, G \cong D$.

6. Type. In virtue of Corollary 5.2 not every finite distributive lattice is representable as the congruence lattice of a 1-unary algebra. However, if the type contains at least two operations then Theorem 5.3, while if it contains an operation which is at least binary then Theorem 2.1 is applicable. Hence we obtain:

6.1. Corollary. *Let us given any type except the* 1*-unary. Then every finite distributive lattice is isomorphic to the congruence lattice of a finite algebra of the given type.*

Acknowledgments. The author is indebted to E. T. Schmidt and E. Fried for helpful comments.

References

[1] G. Birkhoff, *Lattice theory,* Amer. Math. Soc. Colloq. vol. 25 (New York, 1948).

[2] D. P. Egorova, The congruence lattice of a unary algebra, in: *Uporjad. Množestva Rešetki,* vol. 5, Izd. Saratovsk. Univ. (Saratov, 1978); pp. 11—44. (Russian)

[3] R. Freese, Congruence lattices of finitely generated modular lattices, in: *Proc. Lattice Theory Conf.,* Univ. Ulm (Ulm, 1975); pp. 62—70.

[4] G. Grätzer and E. T. Schmidt, On congruence lattices of lattices, *Acta Math. Acad. Sci. Hungar.,* **13** (1962), 179—185.

[5] B. Jónsson, *Topics in universal algebra,* Lecture Notes in Math. 250, Springer (Berlin, 1972).

[6] K. H. Kim and F. W. Roush, Regular rings and distributive lattices, *Commun. Algebra,* **8** (1980), 1283—1290.

[7] R. Kochendörffer, Über treue irreduzible Darstellungen endlicher Gruppen, *Math. Nachr.,* **1** (1948), 25—39.

[8] J. Kuntzmann, Contribution a l'étude des chaines principales d'un groupe fini, *Bull. Sci. Math.* (2), **71** (1947), 155—164.

[9] R. McKenzie, Finite forbidden lattices, in: *Universal Algebra and Lattice Theory* (Proc. Conf. Puebla 1982), Lecture Notes in Math. 1004, Springer (Berlin, 1983); pp. 176—205.

[10] P. P. Pálfy, Unary polynomials in algebras. I, *Algebra Universalis,* **18** (1984), 262—273.

[11] P. P. Pálfy, On partial ordering of chief factors in solvable groups, *Manuscripta Math.*, **55** (1986), 219—232.
[12] P. P. Pálfy and P. Pudlák, Congruence lattices of finite algebras and intervals in subgroup lattices of finite groups, *Algebra Universalis*, **11** (1980), 22—27.
[13] P. Pudlák, Distributivity of strongly representable lattices, *Algebra Universalis*, **7** (1977), 85—92.
[14] R. Quackenbush and B. Wolk, Strong representation of congruence lattices, *Algebra Universalis*, **1** (1971/72), 165—166.
[15] E. T. Schmidt, *Kongruenzrelationen algebraischer Strukturen*, VEB Deutscher Verlag d. Wiss. (Berlin, 1969).
[16] E. T. Schmidt, Every finite distributive lattice is the congruence lattice of some modular lattice, *Algebra Universalis*, **4** (1974), 49—57.
[17] E. T. Schmidt, *A survey on congruence lattice representations*, Teubner (Leipzig, 1982).
[18] H. L. Silcock, Generalized wreath products and the lattice of normal subgroups of a group, *Algebra Universalis*, **7** (1977), 361—372.
[19] J. Tůma, Some finite congruence lattices. I, preprint, 1985.
[20] S. M. Vovsi, Lattices of invariant subspaces of group representations, *Algebra Universalis*, **12** (1981), 221—223.

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
PF. 127
1364 BUDAPEST, HUNGARY

Acta Sci. Math., 51 (1987), 163—183

A partial ordering for the chief factors of a solvable group

G. PAZDERSKI

Dedicated to the memory of András P. Huhn

Let G be a finite solvable group. The chief factors of G can be considered in a natural way as representation modules of G. If M, N are normal subgroups such that M/N is a chief factor of G then the centralizer $C_G(M/N)$ of M/N in G, consisting of all elements $g \in G$ with $x^{-1}g^{-1}xg \in N$ for each $x \in M$, is the kernel of the representation which G takes on M/N. Chief factors M_1/N_1, M_2/N_2 are G-isomorphic, denoted by $M_1/N_1 \cong_G M_2/N_2$, iff they afford equivalent representations of G over the same finite prime field. The class of chief factors of G which are G-isomorphic to the chief factor M/N will be denoted by $[M/N]$. We introduce a partial ordering for the classes of G-isomorphic chief factors of G: The class $[M_1/N_1]$ is said to be greater than the class $[M_2/N_2]$, denoted by $[M_1/N_1] \succ [M_2/N_2]$, if there is a chief factor M_1^*/N_1^* of G such that

$$G \geqq \ldots \geqq M_1^* > N_1^* \geqq \ldots \geqq C_G(M_2/N_2), \quad M_1^*/N_1^* \cong_G M_1/N_1.$$

The set of classes of G-isomorphic chief factors of G together with the partial ordering $\geqq$ will be denoted by $\mathfrak{H}(G)$. This paper deals with some relations between the structure of G and properties of the poset $\mathfrak{H}(G)$.

In Section 1 some basic facts are treated. They concern maximal and minimal elements of $\mathfrak{H}(G)$ in connection with the Fitting subgroup of G, a "colouring" of $\mathfrak{H}(G)$ with the primes dividing $|G|$, and the poset of classes of chief factors belonging to factor groups and direct products. In Section 2 the influence of the partial ordering on pieces of a chief series is investigated. In particular the structure of monotonic pieces of p-chief factors is clarified. In Section 3 the well known concept of a p-series, due to P. Hall and G. Higman [1], is generalized to that of a $\mathfrak{P}$-series of G, where $\mathfrak{P}$ denotes any subset of $\mathfrak{H}(G)$. We are concerned with a comparison between the length $l(\mathfrak{P})$ of $\mathfrak{P}$ and the length $l_{\mathfrak{P}}(G)$ of the $\mathfrak{P}$-series of G. In general $l(\mathfrak{P}) \leqq l_{\mathfrak{P}}(G)$ holds, but in important cases we have equality here. In the concluding Section 4

Received April 24, 1986, and in revised form August 14, 1986.

separating normal subgroups are considered. A normal subgroup N of G is called separating, if no chief factor of G between G and N is G-isomorphic to one below N. This is a generalization of the notion of a normal Hall subgroup. Let $\mathfrak{H}_G(N)$ denote the set of those elements of $\mathfrak{H}(G)$ which are represented by chief factors occurring below N. Then $N \mapsto \mathfrak{H}_G(N)$ is an isomorphism of the lattice $\mathfrak{N}_{\text{sep}}(G)$ of all separating normal subgroups of G into the lattice Low $(\mathfrak{H}(G))$ of all lower segments of the power set Pow $(\mathfrak{H}(G))$ of $\mathfrak{H}(G)$. We characterize those groups, for which this isomorphism is onto.

Notation. All groups considered are finite; G always denotes a solvable group; $H \leqq G$, $H < G$ means: H is a subgroup, proper subgroup of G, respectively; if, in addition, H is normal in G, we write $\trianglelefteqq$, $\lhd$, respectively; analogously $\subseteqq$, $\subset$ denotes inclusion, proper inclusion for sets; $\langle M \rangle :=$ subgroup generated by the subset M of a group; $C_G(M/N) := \{g \mid x^{-1}g^{-1}gx \in N$ for each $x \in M\}$ for normal subgroups M, N of G with $M \geqq N$ is the centralizer of M/N in G; if $U \subseteqq C_G(M/N)$, then M/N is said to be U-central; M/N is central in G means M/N is G-central; a chief factor M/N is said to be situated above or below a normal subgroup K of G, if $M > N \geqq K$ or $K \geqq M > N$, respectively; $[a, b] := a^{-1}b^{-1}ab$; $[H, K] := \langle [a, b] \mid a \in H, b \in K \rangle$; $G' := [G, G]$; p, q denote primes; $F(G) :=$ Fitting subgroup of G, i.e. the maximal nilpotent normal subgroup of G; $F_p(G) := O_{p'p}(G)$, i.e. the maximal p-nilpotent normal subgroup of G; Soc $G :=$ socle of G, i.e. the product of all minimal normal subgroups of G; $A \propto B :=$ semi-direct product of the groups A and B, where B is normal; $l(M) :=$ length of the poset M, i.e. the maximum of the lengths of all chains in M; Pow $M :=$ power set of the set M; $\mathfrak{N}(G) :=$ lattice of all normal subgroups of G. The rest of the notation is introduced in the text in so far as it is not standard (see also [2]).

1. The partially ordered set $\mathfrak{H}(G)$

Firstly we will show that the relation $\geqq$ for the classes of chief factors defined in the introduction is indeed a partial ordering.

1.1. Lemma. *The relation $\geqq$ for the classes of chief factors of a solvable group is reflexive, transitive and antisymmetric.*

Proof. Reflexivity is clear. If $[M_1/N_1] > [M_2/N_2]$ and $[M_2/N_2] > [M_3/N_3]$, then we may assume that $M_1 > N_1 \geqq C_G(M_2/N_2)$ and $M_2 > N_2 \geqq C_G(M_3/N_3)$. But $C_G(M_2/N_2) \geqq M_2$, so $M_1 > N_1 \geqq C_G(M_3/N_3)$ and transitivity is proved. In order to prove antisymmetry assume $[M_1/N_1] > [M_2/N_2]$ and $[M_2/N_2] > [M_1/N_1]$ hold simultaneously. We use the transitivity of $>$, already proved above, and obtain

$[M_1/N_1] > [M_1/N_1]$. This cannot be valid because all chief factors of G, isomorphic to M_1/N_1, are below $C_G(M_1/N_1)$.

The poset of the classes of G-isomorphic chief factors of G will be denoted by $\mathfrak{H}(G)$.

1.2. Examples. (1) If G is a group of prime power order, then $\mathfrak{H}(G)$ has only one element and vice versa.

(2) If G is nilpotent with r different primes in its order, then $\mathfrak{H}(G)$ is an antichain (i.e. the elements of $\mathfrak{H}(G)$ are pairwise incomparable) with r elements and vice versa.

(3) If G is such that $C_G(M/N)=M$ for each chief factor M/N of G, then $\mathfrak{H}(G)$ is a chain, and if in a chief series of G no two chief factors are G-isomorphic, the converse is also true.

(4) Let G be the dihedral group of order $2n$, where the number r of different odd primes in n is not zero. Then $\mathfrak{H}(G)$ has $r+1$ elements, namely a unique maximal one covering the remaining elements, which are minimal.

Obviously $[M_1/N_1] \geqq [M_2/N_2]$ yields $C_G(M_1/N_1) \geqq C_G(M_2/N_2)$. Hence $F(G)$, which is known to be the intersection of the centralizers $C_G(M/N)$ where M/N ranges over all chief factors of G, satisfies

$$F(G) = \bigcap_{\substack{[M/N] \text{ minimal} \\ \text{in } \mathfrak{H}(G)}} C_G(M/N).$$

The maximal and the minimal members of $\mathfrak{H}(G)$ are characterized in the following

1.3. Proposition. (1) *The maximal elements of* $\mathfrak{H}(G)$ *are exactly those represented by the central chief factors of* G.

(2) *The minimal elements of* $\mathfrak{H}(G)$ *are exactly those having no representative above* $F(G)$.

Proof. (1) is trivial.

(2) Assume $[M/N]$ is not a minimal element of $\mathfrak{H}(G)$. Then there are chief factors M_1/N_1, M_2/N_2 of G with $M/N \cong_G M_1/N_1$, $M_1 > N_1 \geqq C_G(M_2/N_2)$. Because of $C_G(M_2/N_2) \geqq F(G)$ it follows that there is a chief factor above $F(G)$, G-isomorphic to M/N. Let, conversely, M/N be a chief factor of G above $F(G)$. Assume $[M/N]$ is minimal in $\mathfrak{H}(G)$. Then for each chief factor M_1/N_1 of G there is no chief factor above $C_G(M_1/N_1)$ which is G-isomorphic to M/N. By the isomorphism theorem it follows that no chief factor G-isomorphic to M/N occurs outside the intersection of all centralizers $C_G(M_1/N_1)$, where M_1/N_1 runs through the chief factors of G. Since this intersection equals $F(G)$, we have a contradiction.

The elements of $\mathfrak{H}(G)$ can be "coloured" by the primes: $[M/N]$ is called a p-element of $\mathfrak{H}(G)$, and will be marked in the graph of $\mathfrak{H}(G)$ by p, if M/N is a p-group. In $\mathfrak{H}(G)$ all p-elements form a subposet $\mathfrak{H}_p(G)$. For a given set π of primes a π-element of $\mathfrak{H}(G)$ is a p-element with any prime $p\in\pi$ and $\mathfrak{H}_\pi(G)$ denotes the subposet of $\mathfrak{H}(G)$ consisting of all π-elements. Obviously $\mathfrak{H}_\pi(G)=\bigcup_{p\in\pi}\mathfrak{H}_p(G)$.

An arbitrary finite group H is said to be π-nilpotent, if it has a normal π'-subgroup N such that H/N is a nilpotent π-group. In any finite group G all π-nilpotent normal subgroups generate a normal subgroup which again is π-nilpotent. This normal subgroup will be designated by $F_\pi(G)$. Clearly, it generalizes the notion of the greatest p-nilpotent normal subgroup $F_p(G)$ of G. We note that

$$F_p(G)=\bigcap_{\substack{[M/N]\text{ minimal}\\ \text{in } \mathfrak{H}_p(G)}} C_G(M/N),$$

$$F_\pi(G)=\bigcap_{\substack{[M/N]\text{ minimal}\\ \text{in } \mathfrak{H}_\pi(G)}} C_G(M/N).$$

Similarly to Proposition 1.3 we have

1.4. Proposition. (1) *The maximal elements of* $\mathfrak{H}_p(G)$ *are exactly those represented by the* $O^{p'}(G)$*-central p-chief factors of* G.

(2) *The minimal elements of* $\mathfrak{H}_p(G)$ *are those elements of* $\mathfrak{H}_p(G)$, *which have no representative above* $F_p(G)$.

1.5. Proposition. (1) *The maximal elements of* $\mathfrak{H}_\pi(G)$ *are exactly those represented by the* $O^{\pi'}(G)$*-central* π*-chief factors of* G.

(2) *The minimal elements of* $\mathfrak{H}_\pi(G)$ *are those elements of* $\mathfrak{H}_\pi(G)$, *which have no representative above* $F_\pi(G)$.

The results stated in Proposition 1.5 will be generalized in Lemma 3.5.

The poset belonging to a direct product can be established from those of the factors in a simple manner.

1.6. Proposition. *Let* $G=G_1\times G_2$. *Then* $\mathfrak{H}(G)$ *arises from* $\mathfrak{H}(G_1)$ *and* $\mathfrak{H}(G_2)$ *by identifying the maximal elements of* $\mathfrak{H}(G_1)$ *and* $\mathfrak{H}(G_2)$ *which are marked with the same prime.*

Proof. The chief factors of G_1 and G_2 can be considered in a natural way as chief factors of G. Let M_1/N_1 be a chief factor of G_1, M_2/N_2 one of G_2 such that $[M_1/N_1]$ and $[M_2/N_2]$ are comparable in $\mathfrak{H}(G)$. If $[M_1/N_1]>[M_2/N_2]$, then there is a chief factor M_1^*/N_1^* of G with

$$M_1^*>N_1^*\geqq C_G(M_2/N_2),\quad M_1^*/N_1^*\cong_G M_1/N_1.$$

Because of $C_G(M_2/N_2)\cong G_1$ we see that M_1/N_1 is G-isomorphic to a chief factor of G_2. Hence M_1/N_1 is centralized by G_1, and since obviously M_1/N_1 is centralized by G_2, we see that M_1/N_1 is both a central chief factor of G_1 and G-isomorphic to a central chief factor of G_2. An analogous consideration works for M_2/N_2 if $[M_1/N_1]<$ $<[M_2/N_2]$. If $[M_1/N_1]=[M_2/N_2]$ in $\mathfrak{H}(G)$, then clearly M_i/N_i is central in G_i $(i=1,2)$ and further M_1/N_1, M_2/N_2 are isomorphic. Conversely, central chief factors of G_1 and G_2 are G-isomorphic if they are isomorphic.

Proposition 1.6 shows that $\mathfrak{H}(G_1\times G_2)$ decomposes into disjoint subposets $\mathfrak{H}(G_1)$, $\mathfrak{H}(G_2)$ if G_1, G_2 have coprime orders. However, $\mathfrak{H}(G)$ can decompose in this way even if G does not decompose directly. An example is the group $SL(2,3)$ (see 1.13 (1) (b)).

We often have to consider the appearance or non-appearance of certain chief factors of G between given normal subgroups of G. For brevity we introduce the following notation: For normal subgroups M, N of G with $M>N$ let $\mathfrak{H}_G(M/N)$ denote the subset of those elements of $\mathfrak{H}(G)$, which have a representative between M and N; further, let $\mathfrak{K}_G(M/N)$ denote the subset of those elements of $\mathfrak{H}(G)$, which do not have representatives above M or below N. We put $\mathfrak{H}_G(M/N)=\mathfrak{K}_G(M/N)=\emptyset$, when $M=N$. Furthermore, $\mathfrak{H}_G(M/1)=:\mathfrak{H}_G(M)$, $\mathfrak{K}_G(M/1)=:\mathfrak{K}(M)$. Obviously,

$$\mathfrak{K}_G(M/N)\subseteqq\mathfrak{H}_G(M/N),\quad \mathfrak{K}_G(M)\subseteqq\mathfrak{H}_G(M).$$

If, for a given subset $\mathfrak{P}\subseteqq\mathfrak{H}(G)$ and normal subgroups M, N of G, $\mathfrak{H}_G(M/N)\subseteqq\mathfrak{P}$ holds, then M/N is said to be a $\mathfrak{P}$-factor of G (or M a normal $\mathfrak{P}$-subgroup in case $N=1$).

If $N\lhd G$ then each chief factor of G lying above N can in a natural way be considered as a chief factor of G/N and vice versa.

1.7. Lemma. (1) *Let $K\lhd G$ and let M/N run through the non-equivalent chief factors of G occurring above K. Then $[M/N]\mapsto[(M/K)/(N/K)]$ is a bijection of $\mathfrak{H}_G(G/K)$ onto $\mathfrak{H}(G/K)$ preserving the partial ordering.*

(2) *$\mathfrak{H}(G/F(G))$ arises from $\mathfrak{H}(G)$ by deleting the minimal elements of $\mathfrak{H}(G)$.*

Proof. (1) is obvious.

(2) Choose $K=F(G)$ in (1) and note that according to Proposition 1.3 (2) $\mathfrak{H}_G(G/K)$ consists of all non-minimal elements of $\mathfrak{H}(G)$.

1.8. Lemma. *$[M_1/N_1]$ is an upper neighbour of $[M/N]$ in $\mathfrak{H}(G)$ iff it has a representative between F and $C_G(M/N)$ but none above F; here F is such that $F/C_G(M/N)$ is the Fitting subgroup of $G/C_G(M/N)$.*

Proof. $[M_1/N_1]>[M/N]$ holds iff M_1/N_1 has a G-isomorphic copy between G and $C_G(M/N)=:K$, and by Lemma 1.7 (1) all those $[M_1/N_1]$ form a subposet of

$\mathfrak{H}(G)$ isomorphic to $\mathfrak{H}(G/K)$. $[M_1/N_1]$ is an upper neighbour of $[M/N]$ iff $[(M_1/K)/(N_1/K)]$ is a minimal element of G/K and by Lemma 1.3 (2) this happens iff there are no chief factors above F/K which are G/K-isomorphic to $(M_1/K)/(N_1/K)$, i.e. iff there is no G-chief factor above F which is G-isomorphic to M_1/N_1.

A condition for the colouring of $\mathfrak{H}(G)$ is in

1.9. Lemma. (1) *Let M_1/N_1 and M/N be chief factors of G such that M_1/N_1 is between F and $C_G(M/N)$, where F is as in Lemma 1.8. Then $|M_1/N_1|$ and $|M/N|$ are relatively prime.*

(2) *In $\mathfrak{H}(G)$ different maximal elements as well as neighbouring elements bear different primes.*

Proof. (1) Let M/N be a p-group, say. Then G/K with $K:=C_G(M/N)$ induces on M/N a faithful irreducible representation over $GF(p)$. Therefore it has no normal p-subgroup $\neq 1$ ([2], p. 485, Satz 5.17). Hence $p\nmid |F: K|$, which yields $p\nmid |M_1/N_1|$.

(2) Since the maximal elements of $\mathfrak{H}(G)$ are the classes of central chief factors, they produce the 1-representations of G. Thus they afford equivalent representations of G, when they have the same prime order. The statement on neighbouring elements comes from (1) in view of Lemma 1.8.

An immediate consequence of the isomorphism theorems is

1.10. Lemma. *Let N_1, N_2 be normal subgroups of G. Then*

(1) $\mathfrak{H}_G(N_1N_2)=\mathfrak{H}_G(N_1)\cup\mathfrak{H}_G(N_2)$.

(2) $\mathfrak{K}_G(N_1\cap N_2)=\mathfrak{K}_G(N_1)\cap\mathfrak{K}_G(N_2)$.

(3) $\mathfrak{H}_G(G/N_1\cap N_2)=\mathfrak{H}_G(G/N_1)\cup\mathfrak{H}_G(G/N_2)$.

(4) $\mathfrak{K}_G(G/N_1N_2)=\mathfrak{K}_G(G/N_1)\cap\mathfrak{K}_G(G/N_2)$.

1.11. Corollary. *Suppose $\mathfrak{H}^*$ and $\mathfrak{K}^*$ are subsets of $\mathfrak{H}(G)$. Then the set*

$$\{N\mid N\trianglelefteq G,\ \mathfrak{H}_G(N)\subseteqq\mathfrak{H}^*,\ \mathfrak{K}_G(N)\supseteqq\mathfrak{K}^*\}, \tag{1.1}$$

if nonvoid, is a lattice, which is an interval of $\mathfrak{N}(G)$.

This enables us to introduce the following normal subgroups. For $\mathfrak{H}^*=\mathfrak{H}(G)$ and arbitrary $\mathfrak{K}^*$ let $N_{\min}(\mathfrak{K}^*)$ be the minimal element of (1.1), and for arbitrary $\mathfrak{H}^*$ and $\mathfrak{K}^*=\emptyset$ let $N_{\max}(\mathfrak{H}^*)$ be the maximal element of (1.1). In particular let $\mathfrak{K}^*$ consists of all minimal elements of $\mathfrak{H}(G)$. Then $N_{\min}(\mathfrak{K}^*)=:N_{\min}(G)$ is a characteristic subgroup of G. On the other hand let $\mathfrak{H}^*$ consist of all non-maximal elements of $\mathfrak{H}(G)$. Then again $N_{\max}(\mathfrak{H}^*)=:N_{\max}(G)$ is a characteristic subgroup of G. $N_{\min}(G)$ is by definition the least normal subgroup of G such that the corresponding factor group has no representative of a minimal element of $\mathfrak{H}(G)$. This characterization from above is accompanied by one from below: namely $N_{\min}(G)$ is also the greatest normal subgroup of G such that each chief factor $N_{\min}(G)/M$ of G represents a minimal element of $\mathfrak{H}(G)$. By Proposition 1.3 (2) $N_{\min}(G) \leqq F(G)$. $N_{\max}(G)$ is by definition the greatest normal subgroup of G which does not include a central chief factor of G. On the other hand, $N_{\max}(G)$ is the least normal subgroup with the property that the corresponding factor group has only central minimal normal subgroups. Obviously, $N_{\max}(G)$ is contained in the nilpotent residual, i.e. the coradical of G with respect to nilpotence. The subgroups $N_{\min}(G)$ and $N_{\max}(G)$ seem not to have been considered yet.

The well-known inclusion $C_G(F(G)) \leqq F(G)$ can be sharpened to

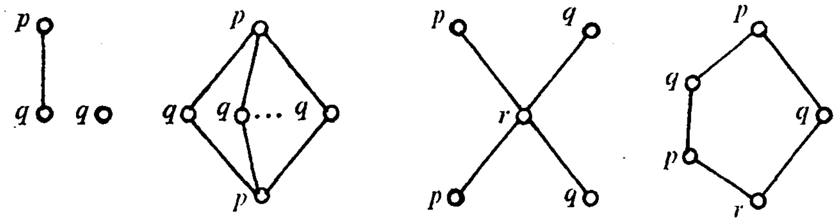
1.12. Proposition. *If N is a normal subgroup of G with $N_{\min}(G) \leqq N \leqq F(G)$, then $C_G(N) \leqq F(G)$.*

Proof. Let $M_1/N_1, \ldots, M_k/N_k$ be representatives below $N_{\min}(G)$ of all minimal elements of $\mathfrak{H}(G)$. If $x \in C_G(N)$, then x centralizes each M_i/N_i $(i=1, \ldots, k)$ and hence $x \in \bigcap_{i=1}^{k} C_G(M_i/N_i) = F(G)$.

We conclude this section with several examples. The following remarks can sometimes be helpful for the construction of a group G with $\mathfrak{H}(G)$ isomorphic to a given prime-coloured poset: If H is a solvable group, which possesses a faithful irreducible representation ∂ over a field $GF(p)$ with representation module M, then the poset $\mathfrak{H}(G)$ of $G = H \propto_{\partial} M$ arises from $\mathfrak{H}(H)$ by adding a new "minimal" element $[M]$ marked by p; that is to say, $\mathfrak{H}(G) = \mathfrak{H}(H) \dot{\cup} [M]$ and $[M]$ as a p-element is covered by all minimal elements of $\mathfrak{H}(H)$. This construction can be generalized by taking pairwise inequivalent faithful irreducible representation modules $M_1, M_2, \ldots, M_k$ of H over arbitrary prime fields. Then for the group $G = H \propto (M_1 \times M_2 \times \ldots \times M_k)$ we get $\mathfrak{H}(G)$ from $\mathfrak{H}(H)$ by adding k new minimal elements $[M_1], [M_2], \ldots, [M_k]$, each of which is covered by every minimal element of $\mathfrak{H}(H)$ and should be marked by the characteristic of the corresponding ground field. These considerations can be modified for the case where the faithful irreducible representations are replaced by irreducible representations with given kernels.

A sufficient condition for a solvable group H to possess a faithful irreducible representation is the existence of an irreducible representation of $\operatorname{Soc} H$, which has a kernel not containing a normal subgroup of G besides 1 (see also [3]). Here the ground field can be arbitrary. This condition is satisfied, if the characteristic of the ground field does not divide $|\operatorname{Soc} H|$ and $\operatorname{Soc} H$ is the direct product of minimal normal subgroups of G no two of which are G-isomorphic.

1.13. Examples. The following posets, coloured with different primes p, q, r are realized by the groups G mentioned below.



(1) (2) (3) (4)

(1) Here two sorts of groups are constructed.

(a) $G=H\times\langle b\rangle$, where H is a minimal non-abelian group of order pq^n with normal Sylow q-subgroup and b has order q.

(b) $G=\langle a\rangle\propto H$, where a has order p and H is an extra-special q-group such that $\langle a\rangle$ acts irreducibly on H/H'. An example in case $p=3$, $q=2$ is the group $G=SL(2,3)=\langle a\rangle\propto H$ with $\operatorname{ord} a=3$ and H the quaternion group. Further groups of this kind of order pq^{2n+1} do exist if $n\geqq 1$, q is odd, and $p\mid q^{2n}-1$, $p\nmid q^i-1$ for $0<i<2n$ (see [6], pp. 14—15).

(2) We start with the cyclic group $\langle a\rangle$ of order p^n and assign to each i with $1\leqq i\leqq n$ an irreducible representation of $\langle a\rangle$ over $GF(q)$ $(q\neq p)$ with kernel $\langle a^{p^i}\rangle$ and corresponding representation module M_i. These representations give rise to a semi-direct product $H=\langle a\rangle\propto(M_1\times M_2\times\ldots\times M_n)$ in which $M_1M_2\ldots M_n$ is the socle. Let $1\neq b_i\in M_i$ for $1\leqq i\leqq n$. Then there is an irreducible representation of the socle over $GF(p)$, which maps $\langle b_1b_2\ldots b_n\rangle$ faithfully. Since its kernel does not contain a normal subgroup of H, there exists a faithful irreducible representation of H over $GF(p)$ with representation module M, say. It produces a semi-direct product $G=H\propto M$, where H acts on M according to this representation. Now $\mathfrak{H}(G)$ has the desired form.

(3) Here we start with the cyclic group $\langle a\rangle$ of order pq and represent it faithfully and irreducibly over $GF(r)$ with representation module M. The associated semi-direct product $H=\langle a\rangle\propto M$ has a faithful irreducible representation over $GF(p)$ as well as over $GF(q)$. The corresponding representation modules M_1, M_2 give rise to a semi-direct product $G=H\propto(M_1\times M_2)$ for which $\mathfrak{H}(G)$ is as desired.

(4) Let $H=\langle a\rangle\propto(M_1\times M_2)$ be the group defined in (2) above with $n=2$. We represent it irreducibly over $GF(p)$ such that M_2 is the kernel. Denote the representation module with M_3 and form the corresponding semi-direct product $K=H\propto M_3$. Obviously $\operatorname{Soc} K=M_2\times M_3$. Hence there is a faithful irreducible representation module M_4 of K over $GF(r)$ and the corresponding semi-direct product $G=K\propto M_4$ has $\mathfrak{H}(G)$ as desired.

2. Monotonic pieces of chief series

2.1. Definition. We will call a piece

$$N_0 > N_1 > \ldots > N_k \tag{2.1}$$

of a chief series of G *monotonic*, if

$$[N_0/N_1] \geqq [N_1/N_2] \geqq \ldots \geqq [N_{k-1}/N_k] \tag{2.2}$$

in the poset $\mathfrak{H}(G)$. If $>$ holds everywhere instead of $\geqq$, then the piece is called *strongly monotonic.*

2.2. Proposition. *G has a strongly monotonic chief series iff G is abnilpotent.*

Recall that a solvable group G is said to be abnilpotent, if $C_G(M/N)=M$ for each chief factor M/N of G (see [7]).

Proof. If G is abnilpotent then there is a chief series (2.1) with $N_0=G$ and $N_k=1$ such that $C_G(N_{i-1}/N_i)=N_{i-1}$ for $i=1, \ldots, k$. This shows the chain to be strongly monotonic. Conversely, let the chief series above be strongly monotonic. Assume there is an index i with $C_G(N_{i-1}/N_i)>N_{i-1}>N_i$. Then in $\mathfrak{H}(G)$ there are less than $i-1$ elements greater than $[N_{i-1}/N_i]$, while otherwise the property of the chain requires that at least $i-1$ members of $\mathfrak{H}(G)$ are greater than $[N_{i-1}/N_i]$. This contradiction yields $C_G(N_{i-1}/N_i)=N_{i-1}$ for each $i=1, \ldots, k$, and therefore G is abnilpotent.

2.3. Proposition. *If G has a monotonic chief series then $\mathfrak{H}(G)$ is a chain.*

Proof. Let (2.1) be a monotonic chief series of G with $N_0=G$, $N_k=1$. When in (2.2) among equal members all but one are deleted, then $\mathfrak{H}(G)$ is seen to be a chain.

The converse of Proposition 2.3 is not true in general. For instance $GL(2, 3)= =S_3 \propto Q$ with Q a quaternion group is a counter-example.

The following proposition gives an insight into the structure of a piece of a chief series in the case when all factor groups are p-groups.

2.4. Proposition. *Let $N_0>N_1>\ldots>N_k$ be a strongly monotonic piece of a chief series of G such that N_0/N_k is a p-group. Then the following hold.*

(1) *N_0/N_k is elementary abelian and $|N_{i-1}\colon N_i|<|N_i\colon N_{i+1}|$ for $i=1, \ldots, k-1$.*

(2) *There is a chain*

$$C_G(N_{k-1}/N_k) = M_1 \geqq M_2 \geqq \ldots \geqq M_k = C_G(N_0/N_k)$$

of normal subgroups M_i of G such that $[M_i, N_j]\leqq N_{i+j}$ for any i, j (here $N_l:=N_k$ for $l\geqq k$) and each factor group M_i/M_{i+1} is abelian with an exponent dividing p.

(3) *If N_0/N_k, as a G-group, is completely reducible, then $C_G(N_{i-1}/N_i)=C_G(N_0/N_i)$ for $i=1, \ldots, k$.*

Proof. (1) To see that the indices $|N_{i-1}\colon N_i|$ increase recall that the p-rank of an irreducible linear group over a field of characteristic p is less than its degree ([5], p. 56, Satz 12). Since N_{i-1}/N_i appears, up to equivalence, above $C_G(N_i/N_{i+1})$, we always have

$$(2.3) \qquad |N_{i-1}\colon N_i| < |N_i\colon N_{i+1}|$$

for $i=1, \dots, k-1$.

In proving the rest of assertion (1) we proceed by induction on k. If $k=1$, there is nothing to prove. Suppose $k>1$ and that the assertion is valid for pieces of chief series which are shorter than the given one. Assume N_0/N_k is non-abelian. Then $N_{k-1}/N_k=(N_0/N_k)'\leqq Z(N_0/N_k)$. Since all $[N_{i-1}/N_i]$ form in $\mathfrak{H}(G)$ a descending chain, we have

$$(2.4) \qquad C_G(N_0/N_1) > C_G(N_1/N_2) > \dots > C_G(N_{k-1}/N_k).$$

Imagining the representation of G on N_0/N_{k-1} and having in mind that by (2.4) $C_G(N_{k-2}/N_{k-1})$ centralizes all N_{i-1}/N_i for $i=1, \dots, k-1$, we find that

$$C_G(N_{k-2}/N_{k-1})/C_G(N_0/N_{k-1})$$

is a p-group. If $C_G(N_{k-1}/N_k)\geqq C_G(N_0/N_{k-1})$, then

$$C_G(N_{k-2}/N_{k-1})/C_G(N_{k-1}/N_k)$$

is a non-trivial normal p-subgroup in the group $G/C_G(N_{k-1}/N_k)$. The latter group has a faithful irreducible representation on N_{k-1}/N_k. This is impossible, and hence there is an $a\in C_G(N_0/N_{k-1})\setminus C_G(N_{k-1}/N_k)$. When acting on N_0/N_k the element a multiplies each element of N_0/N_k with an element of the center of this group, however, it does not fix the commutator subgroup elementwise. This cannot happen. Thus we conclude that N_0/N_k is commutative. By the induction hypothesis N_1/N_k is elementary abelian and so is N_0/N_1. Assume that N_0/N_k is not elementary abelian. Then $\Omega_1(N_0/N_k)=N_1/N_k$ and consequently $|\mho_1(N_0/N_k)|=|N_0\colon N_1|$. Since $\mho_1(N_0/N_k)\leqq N_1/N_k$, the group N_1/N_k has a G-invariant subgroup, the order $|N_0\colon N_1|$ of which is by (2.3) less than the order of each G-chief factor between N_1 and N_k, a contradiction.

(2) For $i=1, \dots, k$ we define

$$M_i := \{g \mid g\in G,\ [g, N_j] \leqq N_{i+j} \text{ for } j = 0, \dots, k-1\}.$$

Then $M_i \trianglelefteq G$ and $M_1\geqq M_2\geqq \dots \geqq M_k$, as can easily be checked. Obviously, $M_1\leqq C_G(N_{k-1}/N_k)$, and by (2.4) $C_G(N_{k-1}/N_k)\leqq M_1$. On the other hand, $M_k\leqq$ $\leqq C_G(N_0/N_k)\leqq M_k$ is clear. By the definition of M_i we have $[M_i, N_j]\leqq N_{i+j}$. As to the commutativity and the exponent of M_i/M_{i+1}, we notice that M_i centralizes N_j/N_{i+j} as well as N_{i+j}/N_{i+j+1}. Therefore in view of (1) $M_i/C_{M_i}(N_j/N_{i+j+1})$ is abelian with exponent dividing p. But $C_{M_i}(N_j/N_{i+j+1})=M_{i+1}$ holds.

(3) If N_0/N_k, as a G-group, is completely reducible, then it is a direct product of subgroups which are admissible with respect to G and G-isomorphic to N_0/N_1, $N_1/N_2, \ldots, N_{k-1}/N_k$. It follows by (2.4) that

$$C_G(N_0/N_i) = \bigcap_{j=1}^{i} C_G(N_{j-1}/N_j) = C_G(N_{i-1}/N_i).$$

The group N_0/N_k mentioned in Proposition 2.4 (3) can be completely reducible or not. Both cases do happen as will be demonstrated in the following

2.5. Example. We start with the group A of order pq^r, where $r = \operatorname{ord} q \bmod p$ and $|A'| = q^r$. It has a faithful irreducible representation ∂ of degree n, say, over $GF(p)$. Let B be the group of all matrices

$$\text{(2.5)} \qquad \begin{pmatrix} 1 & \alpha_1 \; \alpha_2 \ldots \alpha_n \\ 0 & \\ \vdots & \partial(a) \\ 0 & \end{pmatrix}, \quad \alpha_1, \alpha_2, \ldots, \alpha_n \in GF(p), \quad a \in A.$$

Obviously $|B| = pq^r p^n$. B has the set of all matrices with $a=1$ as a normal subgroup N of order p^n; namely it is the kernel of the homomorphism mapping the matrix (2.5) onto the element $a \in A$. We can write $B = A_1 \propto N$ with $A_1 \cong A$. Let N_0 denote a module of dimension $n+1$ over $GF(p)$, on which B acts according to (2.5), and define G to be the appropriate semi-direct product $G := B \propto N_0$. Thus N_0 appears as a normal subgroup of order p^{n+1} of G and contains a normal subgroup N_1 of order p^n of G, on which B acts via ∂. Now G has the chief series

$$G > A_1' N N_0 > N N_0 > N_0 > N_1 > N_2 = 1.$$

We look at the piece $N_0 > N_1 > N_2 = 1$. Since B is represented faithfully on N_0, we have $C_G(N_0) = C_G(N_0/N_2) = N_0$. Further, N_1 is centralized by N_0 and by those elements of B, for which $a=1$ holds in (2.5), i.e. by the elements of N. Hence $C_G(N_1) = C_G(N_1/N_2) = NN_0$, and we obtain $C_G(N_1/N_2) > C_G(N_0/N_2)$. Obviously, N_0/N_1 is a central chief factor of G isomorphic to $G/A_1'NN_0$. Because of $G > A_1'NN_0 > C_G(N_1/N_2)$ we have $[N_0/N_1] > [N_1/N_2]$. So the piece $N_0 > N_1 > N_2 = 1$ is monotonic. By Proposition 2.4 (3) N_0/N_2 does not decompose as a G-group.

To get an example with decomposing factor group N_0/N_2 we can proceed in a similar way. However we take in (2.5) only those matrices with $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Then $N=1$ and the chief series now is

$$G > A_1' N_0 > N_0 > N_1 > N_2 = 1.$$

We have $C_G(N_1/N_2) = C_G(N_0/N_2) = N_0$ and $G/A_1'N_0 \cong_G N_0/N_1$, implying that $[N_0/N_1] > [N_1/N_2]$. Here N_0/N_2 decomposes into two minimal normal subgroups of G, as it is immediately seen by the shape of (2.5) in view of the vanishing α_i.

3. $\mathfrak{P}$-series

One of the central notions in the theory of solvable groups is that of a p-series introduced by P. Hall and G. Higman [1]. Using the poset $\mathfrak{H}(G)$ we can define a more general concept, namely that of a $\mathfrak{P}$-series, where $\mathfrak{P}$ is an arbitrary subset of $\mathfrak{H}(G)$. As in the classical theory, we are interested in the connections between the members of the $\mathfrak{P}$-series and the intersections of the centralizers of chief factors representing elements of $\mathfrak{P}$. Furthermore, the length $l(\mathfrak{P})$ of $\mathfrak{P}$ as a poset is related to the length of the $\mathfrak{P}$-series, and finally Sylow-tower-like theorems are formulated by means of the ordering in $\mathfrak{H}(G)$.

3.1. Definition. Let $\mathfrak{P}\subseteqq\mathfrak{H}(G)$ and put $\mathfrak{P}':=\mathfrak{H}(G)\setminus\mathfrak{P}$. The *upper $\mathfrak{P}$-series* of G is defined as

$$1 = P_0 \leqq Q_0 < P_1 \leqq Q_1 < \ldots < P_l \leqq Q_l = G, \tag{3.1}$$

where Q_i is the greatest normal subgroup of G containing P_i such that $\mathfrak{H}_G(Q_i/P_i)\subseteqq\mathfrak{P}'$ while P_{i+1} is the greatest normal subgroup containing Q_i such that P_{i+1}/Q_i is nilpotent and $\mathfrak{H}_G(P_{i+1}/Q_i)\subseteqq\mathfrak{P}$; this works upwards inductively when starting with $P_0=1$. The number l in (3.1) is called the $\mathfrak{P}$-length of G and will be denoted by $l_{\mathfrak{P}}(G)$.

As was already mentioned, the chain (3.1) coincides with the upper p-series if $\mathfrak{P}=\mathfrak{H}_p(G)$ and it is exactly the upper nilpotent series if $\mathfrak{P}=\mathfrak{H}(G)$. Hence in these cases $l_{\mathfrak{P}}(G)$ is the p-length or the nilpotent length of G, respectively. In case $\mathfrak{P}=\mathfrak{H}_\pi(G)$ the chain (3.1) will be called the upper π-chain and the corresponding length $l_{\mathfrak{P}}(G)$ the π-length of G, denoted by $l_\pi(G)$. The members P_i, Q_i of the chain (3.1) are normal, but need not be characteristic in G. P_1 is called the greatest $\mathfrak{P}$-nilpotent normal subgroup of G and is denoted by $F_{\mathfrak{P}}(G)$.

The upper $\mathfrak{P}$-length is in a certain sense minimal. This is shown in

3.2. Lemma. *Let $\mathfrak{P}\subseteqq\mathfrak{H}(G)$, $\mathfrak{P}':=\mathfrak{H}(G)\setminus\mathfrak{P}$ and let*

$$1 = P_0^* \leqq Q_0^* \leqq P_1^* \leqq Q_1^* \leqq \ldots \leqq P_k^* \leqq Q_k^* = G$$

be a chain of normal subgroups of G such that $\mathfrak{H}_G(Q_i^*/P_i^*)\subseteqq\mathfrak{P}'$ $(i=0, \ldots, k)$ and P_{i+1}^*/Q_i^* is nilpotent with $\mathfrak{H}_G(P_{i+1}^*/Q_i^*)\subseteqq\mathfrak{P}$ $(i=0, \ldots, k-1)$. Then, for the members P_i, Q_i of the chain (3.1) we have

$$P_i^* \leqq P_i, \quad Q_i^* \leqq Q_i \quad \textit{for all } i.$$

In particular, $l_{\mathfrak{P}}(G)\leqq k$.

Proof. $P_0^*\leqq P_0$ holds trivially. Let $P_i^*\leqq P_i$ for a certain i. Then $Q_i^*\leqq Q_i$ can be proved in the following way. Since

$$Q_i^*P_i/P_i \cong_G Q_i^*/Q_i^*\cap P_i \quad\text{and}\quad P_i^* \leqq Q_i^*\cap P_i,$$

the factor group $Q_i^* P_i/P_i$ includes only chief factors of G representing elements of $\mathfrak{P}'$. So $Q_i^* P_i/P_i \leqq Q_i/P_i$ and hence $Q_i^* \leqq Q_i$ holds. In a similar manner we get $P_{i+1}^* \leqq P_{i+1}$. The assertion now follows by induction.

3.3. Corollary. *For* $\mathfrak{P} \subseteqq \mathfrak{H}(G)$ *let*

$$[M_1/N_1] < [M_2/N_2] < \ldots < [M_r/N_r]$$

be a chain in $\mathfrak{P}$. *For each* $i=1, \ldots, r$ *let* $j(i)$ *be the maximal* j *with* $[M_i/N_i] \in \mathfrak{H}_G(P_j/Q_{j-1})$, *where the meaning of* P_j, Q_{j-1} *is as in* (3.1). *Then*

(1) $1 \leqq j(1) < j(2) < \ldots < j(r) \leqq l_{\mathfrak{P}}(G)$.

(2) $l(\mathfrak{P}) \leqq l_{\mathfrak{P}}(G)$.

Proof. (1) Since $P_{j(i)}/Q_{j(i)-1}$ is nilpotent, $P_{j(i)} \leqq C_G(M_i/N_i)$. This yields, in view of $[M_i/N_i] < [M_{i+1}/N_{i+1}]$, that there occurs a chief factor above $P_{j(i)}$ which is G-isomorphic to M_{i+1}/N_{i+1}, hence $j(i) < j(i+1)$.

(2) is a consequence of (1).

3.4. Remark. It can happen that $l(\mathfrak{P}) < l_{\mathfrak{P}}(G)$. Take for instance $G := GL(2,3) = S_3 \propto H$, where S_3 denotes the symmetric group of degree 3 and H denotes the quaternion group, and let $\mathfrak{P}$ consist of the unique class of central chief factors of G. Then $l(\mathfrak{P})=1$ and the $\mathfrak{P}$-series of G has $P_0=1$, $Q_0=1$, $P_1=Z(G)$, $Q_1=A_3 \propto H$, $P_2=G$, $Q_2=G$ with A_3 the alternating group of degree 3 in S_3. Hence $l_{\mathfrak{P}}(G)=2$.

Next we are looking for conditions, which guarantee that $l(\mathfrak{P})=l_{\mathfrak{P}}(G)$ holds instead of $l(\mathfrak{P}) \leqq l_{\mathfrak{P}}(G)$.

For a given subset $\mathfrak{P} \subseteqq \mathfrak{H}(G)$ we define

$$O_{\mathfrak{P}}(G) := \langle N \mid N \trianglelefteq G, \mathfrak{H}_G(N) \subseteqq \mathfrak{P} \rangle,$$

$$O^{\mathfrak{P}}(G) := \bigcap_{\substack{N \trianglelefteq G \\ \mathfrak{H}_G(G/N) \subseteqq \mathfrak{P}}} N,$$

$$C_G(\mathfrak{P}) := \bigcap_{[M/N] \in \mathfrak{P}} C_G(M/N).$$

If $\mathfrak{P}=\emptyset$, then by the definition of $\mathfrak{H}_G$ and $\mathfrak{K}_G$ we have $O_{\mathfrak{P}}(G)=1$, $O^{\mathfrak{P}}(G)=G$; additionally we define in this case $C_G(\mathfrak{P}) := G$.

The groups $O_{\mathfrak{P}}(G)$, $O^{\mathfrak{P}}(G)$ are generalizations of the characteristic subgroups $O_\pi(G)$, $O^\pi(G)$ for a set of primes π; they appear with $\mathfrak{P}=\mathfrak{H}_\pi(G)$. They are also connected with the groups $N_{\min}$ and $N_{\max}$ introduced in Section 1, namely

$$N_{\min}(\mathfrak{P}) = O^{\mathfrak{P}'}(G), \quad N_{\max}(\mathfrak{P}) = O_{\mathfrak{P}}(G).$$

As a common generalization of Propositions 1.3, 1.4 and 1.5 we obtain

3.5. Lemma. *An element* $[M/N]$ *of* $\mathfrak{H}(G)$ *is not surpassed by any element of* $\mathfrak{P}$ *iff* M/N *is* $O^{\mathfrak{P}'}(G)$*-central, and it does not surpass any element of* $\mathfrak{P}$ *iff* $[M/N]\in\mathfrak{K}_G(C_G(\mathfrak{P}))$.

Proof. $[M/N]$ is not surpassed by any element of $\mathfrak{P}$ iff $\mathfrak{H}_G(G/C_G(M/N))\subseteq\mathfrak{P}'$ and this happens iff $C_G(M/N)\supseteq O^{\mathfrak{P}'}(G)$. On the other hand, $[M/N]$ does not surpass any element of $\mathfrak{P}$ iff for each $[M_1/N_1]\in\mathfrak{P}$ we have $[M/N]\notin\mathfrak{H}_G(G/C_G(M_1/N_1))$, which means $[M/N]\in\mathfrak{K}_G(C_G(M_1/N_1))$. By Lemma 1.10 (2) this is equivalent to $[M/N]\in\mathfrak{K}_G(C_G(\mathfrak{P}))$.

We define an ascending centralizer chain with respect to a given non-empty subset $\mathfrak{P}\subseteq\mathfrak{H}(G)$. Let

$$\emptyset=\mathfrak{P}_0\subset\mathfrak{P}_1\subset\ldots\subset\mathfrak{P}_l=\mathfrak{P}$$

be the unique chain of subsets of $\mathfrak{P}$ such that $\mathfrak{P}_i\setminus\mathfrak{P}_{i-1}$ consists of the minimal elements of $\mathfrak{P}\setminus\mathfrak{P}_{i-1}$ $(i=1,\ldots,l)$. Obviously, l coincides with the length $l(\mathfrak{P})$ of $\mathfrak{P}$. The equality $C_G(\mathfrak{P}\setminus\mathfrak{P}_{i-1})=C_G(\mathfrak{P}_i\setminus\mathfrak{P}_{i-1})$ is easily seen, too.

3.6. Lemma. *We have, for* $\emptyset\neq\mathfrak{P}\subseteq\mathfrak{H}(G)$,

(1) $1<C_G(\mathfrak{P}_1\setminus\mathfrak{P}_0)<C_G(\mathfrak{P}_2\setminus\mathfrak{P}_1)<\ldots<C_G(\mathfrak{P}_l\setminus\mathfrak{P}_{l-1})\leqq G$.
Here the chief factors of G *between* $C_G(\mathfrak{P}_l\setminus\mathfrak{P}_{l-1})$ *and* G *represent only elements of* $\mathfrak{P}'$, *i.e.,* $\mathfrak{H}_G(G/C_G(\mathfrak{P}_l\setminus\mathfrak{P}_{l-1}))\subseteq\mathfrak{P}'$.

(2) $P_i\leqq C_G(\mathfrak{P}_i\setminus\mathfrak{P}_{i-1})$ for $i=1,\ldots,l$, *where the* P_i *come from the chain in* (3.1) *related to* $\mathfrak{P}$.

Proof. (1) Since $C_G(\mathfrak{P}_1\setminus\mathfrak{P}_0)$ contains the Fitting subgroup of G, it differs from 1. Next we prove that $C_G(\mathfrak{P}_i\setminus\mathfrak{P}_{i-1})\neq C_G(\mathfrak{P}_{i+1}\setminus\mathfrak{P}_i)$. By Lemma 3.5 $\mathfrak{P}_i=\mathfrak{P}\cap\mathfrak{K}_G(C_G(\mathfrak{P}_i\setminus\mathfrak{P}_{i-1}))$. Here on the left hand side with different values of i always different sets $\mathfrak{P}_i$ occur. Hence the corresponding groups $C_G(\mathfrak{P}_i\setminus\mathfrak{P}_{i-1})$ must be different. If $C_G(\mathfrak{P}_l\setminus\mathfrak{P}_{l-1})\leqq N<M\leqq G$, where M/N is a chief factor of G, then $[M/N]\notin\mathfrak{K}_G(C_G(\mathfrak{P}_l\setminus\mathfrak{P}_{l-1}))\supseteq\mathfrak{P}_l=\mathfrak{P}$. Hence we obtain $[M/N]\notin\mathfrak{P}$.

(2) Assume $P_i\not\leqq C_G(\mathfrak{P}_i\setminus\mathfrak{P}_{i-1})$ for some $i\geqq 1$ and take i minimal. There is an $[M/N]\in\mathfrak{P}_i\setminus\mathfrak{P}_{i-1}$ which is not centralized by P_i. Therefore $[M/N]\notin\mathfrak{H}_G(G/P_i)$. Because of the nilpotence of P_i/Q_{i-1}, we have $[M/N]\notin\mathfrak{H}_G(P_i/Q_{i-1})$. Consequently $i\geqq 2$ and $[M/N]\in\mathfrak{K}_G(Q_{i-1})$. Since $[M/N]\in\mathfrak{P}$, it follows that $[M/N]\in\mathfrak{K}_G(P_{i-1})$. By the minimality of i, $P_{i-1}\leqq C_G(\mathfrak{P}_{i-1}\setminus\mathfrak{P}_{i-2})$. Now we get that $[M/N]$ does not surpass an element of $\mathfrak{P}_{i-1}\setminus\mathfrak{P}_{i-2}$, a contradiction.

Lemma 3.6 (2) yields, for $i=1$, the following

3.7. Corollary. $F_{\mathfrak{P}}(G)\leqq C_G(\mathfrak{P})$ *for* $\emptyset\neq\mathfrak{P}\subseteq\mathfrak{H}(G)$.

3.8. Corollary. $l(\mathfrak{P})=l_{\mathfrak{P}}(G)$ *holds if in Lemma* 3.6 (2) $P_l=C_G(\mathfrak{P}\setminus\mathfrak{P}_{l-1})$ *with* $l=l(\mathfrak{P})$.

Proof. By Lemma 3.6 (1) the assumption yields that $\mathfrak{H}(G/P_l)\subseteq\mathfrak{P}'$ and therefore $Q_l=G$. Hence, in view of Lemma 3.2, $l_{\mathfrak{P}}(G)\leqq l(\mathfrak{P})$, and by Corollary 3.3 (2), $l(\mathfrak{P})=l_{\mathfrak{P}}(G)$.

3.9. Remark. In Lemma 3.6 (2) equality as well as inequality can happen. The case of equality is discussed in the next theorem. Inequality holds for the group $G:=SL(2,3)=\langle a\rangle\propto H$ which was already mentioned in Example 1.13 (1) (b), if we take $\mathfrak{P}:=\{[Z(H)/1]\}$. Then $P_1=F_{\mathfrak{P}}(G)=Z(H)$, $C_G(\mathfrak{P}_1\setminus\mathfrak{P}_0)=C_G(\mathfrak{P})=G$. This shows also that the converse of Corollary 3.8 need not be true, since $l(\mathfrak{P})=l_{\mathfrak{P}}(G)$.

3.10. Theorem. *Suppose in Lemma 3.6 (2) we have $C_G(\mathfrak{P}_i\setminus\mathfrak{P}_{i-1})>P_i$ for a certain index $i\geqq 1$, and choose i minimal with this property. Let T/P_i be the product of all minimal normal subgroups of G/P_i contained in $C_G(\mathfrak{P}_i\setminus\mathfrak{P}_{i-1})/P_i$. Then T/Q_{i-1} is nilpotent and $\mathfrak{H}_G(T/P_i)\subseteq\mathfrak{P}'$, $\mathfrak{H}_G(P_i/Q_{i-1})\subseteq\mathfrak{P}\setminus\mathfrak{P}_{i-1}$. Furthermore, to each element $[L_1/M_1]\in\mathfrak{H}_G(T/P_i)$ there exists a piece $L>M>N$ of a chief series of G such that the following conditions* (1) *through* (4) *are satisfied:*

(1) $T\geqq L,\quad P_i\geqq M>N\geqq Q_{i-1}$.
(2) $[L/M]=[L_1/M_1]$.
(3) *L/N has prime power order.*
(4) *L/N is indecomposable as a G-group.*

Proof. By the definition of i we have $C_G(\mathfrak{P}_j\setminus\mathfrak{P}_{j-1})=P_j$ for $j=1,\dots,i-1$. Hence, in view of Lemma 3.5, $\mathfrak{P}_j\setminus\mathfrak{P}_{j-1}\subseteq\mathfrak{K}_G(P_j)$ for $j=1,\dots,i-1$ and so $\mathfrak{P}_{i-1}\subseteq\mathfrak{K}_G(P_{i-1})$. This yields $\mathfrak{H}_G(P_i/Q_{i-1})\subseteq\mathfrak{P}\setminus\mathfrak{P}_{i-1}$. Hence T centralizes all G-chief factors between P_i and Q_{i-1}. Those between T and P_i are centralized by T, too. It follows that T/Q_{i-1} centralizes all of its own chief factors, and therefore T/Q_{i-1} is nilpotent. By the definition of P_i no chief factor of G between T and P_i can belong to $\mathfrak{P}$. So $\mathfrak{H}_G(T/P_i)\subseteq\mathfrak{P}'$.

Assume now that to a given L_1/M_1 no piece $L>M>N$ with the properties (1) through (4) exists. We can find a piece $K_1>F_1>F_2$ of a chief series of G with

$$T\geqq K_1>F_1>F_2\geqq Q_{i-1},\quad F_1=P_i,\quad L_1/M_1\cong_G K_1/F_1.$$

Since $[F_1/F_2]\in\mathfrak{P}\setminus\mathfrak{P}_{i-1}$, K_1 centralizes F_1/F_2.

If $|K_1/F_1|$ is coprime to $|F_1/F_2|$, then by Schur—Zassenhaus' Theorem there exists a normal subgroup K_2 of G such that $K_1>K_2>F_2$ and $K_1/F_1\cong_G K_2/F_2$, $F_1/F_2\cong_G K_1/K_2$.

If K_1/F_2 has prime power order, then by assumption K_1/F_2 decomposes as a G-group completely. Hence, again there is a normal subgroup K_2 of G with $K_1>$ $>K_2>F_2$, $K_1/F_1\cong_G K_2/F_2$, $F_1/F_2\cong_G K_1/K_2$.

If $F_2>Q_{i-1}$, take a new chief factor F_2/F_3 of G with $F_3\geqq Q_{i-1}$. Considering the chain $K_2>F_2>F_3$ we can find, as above, a normal subgroup K_3 of G with $K_2>$

$>K_3>F_3$, $K_2/F_2 \cong_G K_3/F_3$, $F_2/F_3 \cong_G K_2/K_3$. Continuing in this way we find chains

$$K_1 > K_2 > \ldots > K_t, \quad P_i = F_1 > F_2 > \ldots > F_t = Q_{i-1}$$

such that always $K_{i-1}/F_{i-1} \cong_G K_i/F_i$. Hence $K_1/F_1 \cong_G K_t/F_t = K_t/Q_{i-1}$. However, $[K_1/F_1] = [L_1/M_1] \notin \mathfrak{P}$, contradicting the construction of Q_{i-1}.

An application of Theorem 3.10 is

3.11. Theorem. *With the notation of Lemma* 3.6 *we have* $C_G(\mathfrak{P}_i \setminus \mathfrak{P}_{i-1}) = P_i$ *for* $i=1, \ldots, l(\mathfrak{P})$, *and hence also* $l(\mathfrak{P}) = l_{\mathfrak{P}}(G)$, *in each of the following cases:*

(1) $\mathfrak{P} = \mathfrak{H}_p(G)$.
(2) $\mathfrak{P} = \mathfrak{H}_\pi(G)$.
(3) $\mathfrak{P} = \mathfrak{H}(G)$.
(4) $\mathfrak{P}$ *is such that* $C_G(M_1/N_1) = M_1$ *holds for every* $[M_1/N_1] \in \mathfrak{P}$.
(5) $\mathfrak{P}$ *is arbitrary and in each chief series of* G *neighbouring factors have coprime orders.*

Proof. (1) and (3) are special cases of (2). Conditions (2), (4), (5) lead to $C_G(\mathfrak{P}_i \setminus \mathfrak{P}_{i-1}) = P_i$ for $i=1, \ldots, l(\mathfrak{P})$. Assume not, and choose i minimal with $C_G(\mathfrak{P}_i \setminus \mathfrak{P}_{i-1}) > P_i$. We utilize Theorem 3.10. The existence of a series $L>M>N$ described there is impossible, because L/N has prime power order and

in case (2) $[L/M] \in \mathfrak{H}_{\pi'}(G)$ and $[M/N] \in \mathfrak{H}_\pi(G)$ hold simultaneously;
in case (4) $M/N \leqq Z(L/N)$, whence $L \leqq C_G(M/N) = M$;
in case (5) $|L/M|$ and $|M/N|$ are relatively prime.

As a consequence of Theorem 3.11 (1)—(3) we have

3.12. Corollary. *The π-length, p-length, and nilpotent length of* G *coincides with the length of* $\mathfrak{H}_\pi(G)$, $\mathfrak{H}_p(G)$, $\mathfrak{H}(G)$, *respectively.*

In particular, we have

3.13. Corollary. $l_\pi(G)=1$ *iff* $\mathfrak{H}_\pi(G)$ *is an antichain,* $l_p(G)=1$ *iff* $\mathfrak{H}_p(G)$ *is an antichain.* G *is nilpotent iff* $\mathfrak{H}(G)$ *is an antichain.*

We close this section with a Sylow-tower-like property.

3.14. Theorem. *Let* $\pi(G) = \pi_1 \dot{\cup} \pi_2 \dot{\cup} \ldots \dot{\cup} \pi_r$ *be a partition of* $\pi(G)$ *into non-empty subsets* π_i. *Suppose an element of* $\mathfrak{H}_{\pi_i}(G)$ *surpasses an element of* $\mathfrak{H}_{\pi_j}(G)$ *only if* $i<j$. *Then* G *has a chain*

$$G = N_0 > N_1 > \ldots > N_r = 1$$

of normal subgroups in which N_{i-1}/N_i *is a nilpotent* π_i*-group, and vice versa.*

Proof. Let $\mathfrak{H}(G)$ satisfy the assumption. We start with $N_0:=G$ and suppose that $N_0, N_1, \ldots, N_{i-1}$ have been constructed for a certain $i \geqq 1$ as normal Hall subgroups of G such that N_{j-1}/N_j is a nilpotent π_j-group for $j=1, \ldots, i-1$. We consider the $\mathfrak{P}$-series of G with $\mathfrak{P}=\mathfrak{H}_{\pi_i}(G)$. Since $\mathfrak{P}$ is an antichain, $l(\mathfrak{P})=1$ and hence by Theorem 3.11 the $\mathfrak{P}$-series is shaped like this:

$$1 = P_0 \leqq Q_0 < P_1 \leqq Q_1 = G.$$

Here P_1/Q_0 is a π_i-group, while Q_0/P_0 and Q_1/P_1 are π_i'-groups. Because of the assumption on the partial ordering in $\mathfrak{H}(G)$, for each π_i-chief factor M/N the factor group $G/C_G(M/N)$ is a $(\pi_1 \dot{\cup} \ldots \dot{\cup} \pi_{i-1})$-group. By Theorem 3.11

$$\bigcap_{[M/N]\in\mathfrak{P}} C_G(M/N) = C_G(\mathfrak{P}) = C_G(\mathfrak{P}_1 \setminus \mathfrak{P}_0) = P_1,$$

so that G/P_1 is a $(\pi_1 \dot{\cup} \ldots \dot{\cup} \pi_{i-1})$-group and therefore $N_{i-1} \leqq P_1$. Put $N_i := N_{i-1} \cap Q_0$. Then $N_{i-1}/N_i \cong_G N_{i-1}Q_0/Q_0 \leqq P_1/Q_0$, so N_{i-1}/N_i is a nilpotent π_i-group. Further, N_i being a subgroup of Q_0 is a π_i'-group. Because of $N_i \leqq N_j$ for $j=1, \ldots, i-1$, N_i is also a π_j'-group for $j=1, \ldots, i-1$. Hence N_i is a Hall subgroup. The converse is trivial.

3.15. Corollary. *G has a Sylow tower belonging to the ordering $p_1 \succ p_2 \succ \ldots \ldots \succ p_r$ of $\pi(G)$ iff $\mathfrak{H}(G)$ is such that $[M_1/N_1] > [M_2/N_2]$, M_1/N_1 a p_i-group and M_2/N_2 a p_j-group, imply $p_i \succ p_j$.*

4. Separating normal subgroups

4.1. Definition. A normal subgroup N of a solvable group G is said to be *separating*, if no chief factor of G has a G-isomorphic copy above N as well as below N.

For instance each normal Hall subgroup is separating. For a normal subgroup N of G the property of being separating is characterized by each of the equations $\mathfrak{H}_G(N)=\mathfrak{K}_G(N)$, $\mathfrak{H}_G(G/N)=\mathfrak{K}_G(G/N)$, $\mathfrak{H}(G)=\mathfrak{H}_G(G/N) \dot{\cup} \mathfrak{H}_G(N)$. All separating normal subgroups of G constitute a sublattice $\mathfrak{N}_{sep}(G)$ of the lattice $\mathfrak{N}(G)$ of all normal subgroups of G.

Recall the notion of a lower segment in a poset M. It is defined as a subset of M, which contains with each of its elements x all elements of M which are surpassed by x. Upper segments are defined analogously.

4.2. Proposition. *If $N \in \mathfrak{N}_{sep}(G)$ then $\mathfrak{H}_G(N)$ is a lower and consequently $\mathfrak{H}_G(G/N)$ is an upper segment of $\mathfrak{H}(G)$.*

Proof. Let $[K/L]\notin \mathfrak{H}_G(N)$ and assume $K \leqq LN$. Then K/L is a chief factor between LN and L, which is G-isomorphic to one between N and $N \cap L$. But then $[K/L]\in \mathfrak{H}_G(N)$. Hence $K \nleqq LN$ and so $LN \cap K = L$. This yields that K/L and LN/L centralize mutually in G/L and so $N \leqq C_G(K/L)$, i.e. each M_1/N_1 with $[M_1/N_1] > > [K/L]$ fulfils $[M_1/N_1] \notin \mathfrak{H}_G(N)$.

All lower segments of $\mathfrak{H}(G)$ form a sublattice of the power set Pow $\mathfrak{H}(G)$; we denote it by Low $\mathfrak{H}(G)$.

4.3. Lemma. *$N \mapsto \mathfrak{H}_G(N)$ $(N \in \mathfrak{N}_{\mathrm{sep}}(G))$ is a lattice isomorphism of $\mathfrak{N}_{\mathrm{sep}}(G)$ into* Low $\mathfrak{H}(G)$. *In particular, $\mathfrak{N}_{\mathrm{sep}}(G)$ is a distributive lattice.*

Proof. Assume there are $N_1, N_2 \in \mathfrak{N}_{\mathrm{sep}}(G)$ with $N_1 \neq N_2$ and $\mathfrak{H}_G(N_1) = \mathfrak{H}_G(N_2)$. Then $N_1 \neq 1$ and $N_2 \neq 1$ and $N_1N_2 > N_2 > 1$. Any G-chief factor between N_1N_2 and N_2 represents an element of $\mathfrak{H}_G(N_1)$. Because of $\mathfrak{H}_G(N_1) = \mathfrak{H}_G(N_2)$ it must also have a copy below N_2, which is impossible. Thus the mapping $N \mapsto \mathfrak{H}_G(N)$ under consideration is injective. From Lemma 1.10 (1) (2) we obtain that it preserves union and intersection. Now, since $\mathfrak{N}_{\mathrm{sep}}(G)$ is isomorphic to a sublattice of the distributive lattice Pow $\mathfrak{H}(G)$, it is distributive, too.

4.4. Corollary. *If $\mathfrak{H}(G)$ is a chain, then $\mathfrak{N}_{\mathrm{sep}}(G)$ is a chain.*

We will characterize those groups G for which the mapping in Lemma 4.3 is onto Low $\mathfrak{H}(G)$. A key role is played by the homogeneous socle of a group introduced below.

4.5. Definition. A normal subgroup N of G is said to be *homogeneous*, if all chief factors of G below N are G-isomorphic. The product of all homogeneous normal subgroups is called the *homogeneous socle* of G and will be denoted by Hos G.

It is easy to show that Hos G is the direct product of maximal homogeneous normal subgroups of G. Since each homogeneous normal subgroup has prime power order and each minimal normal subgroup is homogeneous, we have

$$F(G) \geqq \mathrm{Hos}\, G \geqq \mathrm{Soc}\, G.$$

Each normal subgroup of G contained in Hos G is a direct product of homogeneous normal subgroups.

The notion of homogeneous factor groups and the homogeneous cosocle can be defined in an analogous manner; but we do not need them in the present paper.

4.6. Theorem. *The following statements are equivalent:*

(1) *$\mathfrak{N}_{\mathrm{sep}}(G)$ is isomorphic to* Low $\mathfrak{H}(G)$, *the lattice of all lower segments of* Pow $\mathfrak{H}(G)$.

(2) *Whenever* $L > M > N$ *is a piece of a chief series of* G, *then either* L/N *decomposes completely as a G-group or* $[L/M] \geqq [M/N]$ *in* $\mathfrak{H}(G)$.

(3) $N_{\min}(G/N) \leqq \mathrm{Hos}\,(G/N)$ *holds for each normal subgroup* N *of* G. (For the definition of $N_{\min}$ see Section 1.)

Proof. (1)$\Rightarrow$(2). Let $L > M > N$ be a piece of a chief series of G such that L/N as a G-group does not decompose completely and L/M, M/N are not G-isomorphic. Assume $[L/M] \ngtr [M/N]$. Let $\mathfrak{K}$ be the set of all $[M_1/N_1] \in \mathfrak{H}(G)$ with $[M_1/N_1] \leqq [L/M]$. Then $\mathfrak{K}$ is a lower segment of $\mathfrak{H}(G)$ and by assumption there is a separating normal subgroup K of G such that $\mathfrak{H}_G(K) = \mathfrak{K}$. We consider the series

$$KL \underset{1}{\geqq} KN \underset{2}{\geqq} K \underset{3}{\geqq} K \cap L \underset{4}{\geqq} K \cap N, \tag{4.1}$$

$$KL \underset{3}{\geqq} L \underset{1}{\geqq} KN \cap L \geqq (K \cap L)N \underset{4}{\geqq} N \underset{2}{\geqq} K \cap N, \tag{4.2}$$

where equally numbered factors are G-isomorphic. This yields the equality $KN \cap L = (K \cap L)N$. In (4.2) we look at the piece $L \geqq (K \cap L)N \geqq N$ and realize, in addition, that up to G-isomorphy L/M does not occur above K and M/N does not occur below K. We obtain in view of (4.1) that L/M cannot be G-isomorphic to a chief factor of G between L and $(K \cap L)N$; further, M/N cannot be G-isomorphic to one between $(K \cap L)N$ and N. It follows

$$L/M \cong_G (K \cap L)N/N, \quad M/N \cong_G L/(K \cap L)N.$$

This leads to

$$L/N \cong_G (L/M) \times (M/N),$$

i.e. L/N decomposes into two minimal normal G-subgroups, contrary to our assumption.

(2) $\Rightarrow$ (3). Suppose neighbouring chief factors of G always have the property described in (2). Let $[M_1/N_1]$ be a minimal element of $\mathfrak{H}(G)$ and let M be a normal subgroup of G such that there is no chief factor between G and M which is G-isomorphic to M_1/N_1; moreover let us choose M minimal with this property. We will show that M is homogeneous in G. Assume not, then there is a chief factor M_2/N_2 of G with $M > M_2 > N_2$ such that between M and M_2 there are only chief factors which are G-isomorphic to M_1/N_1 and $M_2/N_2 \ncong_G M_1/N_1$. In particular, M/M_2 is a group of prime power order. Let us choose M_2, N_2 such that M/M_2 is as small as possible. If $M \not\leqq C_G(M_2/N_2)$, then $[M_1/N_1] > [M_2/N_2]$, so $[M_1/N_1]$ would not be minimal. Hence $M \leqq C_G(M_2/N_2)$. If the orders of M/M_2 and M_2/N_2 were relatively prime, then $M/N_2 \cong_G (M/M_2) \times (M_2/N_2)$ and there would exist a normal subgroup K of G with $M/K \cong_G M_2/N_2$, contradicting the choice of M. Hence M/N_2 has prime power order. There exists L with $M \geqq L > M_2 > N_2$ such that L/M_2 and M_2/N_2 are chief factors of G. Since $[L/M_2] = [M_1/N_1] \not\leqq [M_2/N_2]$, by assumption L/N_2 de-

composes completely. But then L/N_2 has a factor group which is G-isomorphic to M_2/N_2, contradicting the choice of M_2 and N_2. This proves that M is homogeneous.

Let $\mathfrak{M}$ be the set of minimal elements of $\mathfrak{H}(G)$. As was shown above, there exists to each element $[M_0/N_0]$ of $\mathfrak{M}$ a homogeneous separating normal subgroup of G involving only chief factors of G which are G-isomorphic to M_0/N_0. The product of these homogeneous normal subgroups belonging to all elements of $\mathfrak{M}$ coincides with $N_{\min}(G)$ and is contained in Hos G. Thus the assertion (3) is proved for $N=1$. Since the assumption (2) is hereditary to factor groups, so is the assertion (3), and we are done.

(3)$\Rightarrow$(1). The assertion that each lower segment $\mathfrak{L}$ of $\mathfrak{H}(G)$ belongs to a separating normal subgroup L of G such that $\mathfrak{H}_G(L)=\mathfrak{L}$ is proved by induction on $|G|$. Suppose $|G|>1$ because in case $|G|=1$ there is nothing to prove. Let $\mathfrak{L}_0$ be the set of minimal elements of $\mathfrak{L}$. Assumption (3) yields that to each element $[M_0/N_0]$ of $\mathfrak{L}_0$ there exists a homogeneous separating normal subgroup of G involving only chief factors which are G-isomorphic to M_0/N_0. Let L_0 be the product of these homogeneous normal subgroups belonging to all elements of $\mathfrak{L}_0$. The factor group G/L_0 has $\mathfrak{L}\setminus\mathfrak{L}_0$ as a lower segment in a natural way. By the induction hypothesis G/L_0 has a separating normal subgroup L/L_0 with $\mathfrak{H}_{G/L_0}(L/L_0)=\mathfrak{L}\setminus\mathfrak{L}_0$. Now L is separating in G and satisfies $\mathfrak{H}_G(L)=\mathfrak{L}$.

We will formulate some conditions, which are sufficient for the properties mentioned in Theorem 4.6.

4.7. Theorem. *The following conditions are equivalent and they imply the properties described in Theorem* 4.6:

(1) *Whenever* $L>M>N$ *is a piece of a chief series of* G *such that* L/N *has prime power order and* $L/M\not\cong_G M/N$, *then* L/N *as a* G*-group decomposes completely.*

(2) $F(G/N)=\mathrm{Hos}\,(G/N)$ *holds for each normal subgroup* N *of* G.

Proof. (1)$\Rightarrow$(2). Assume $F(G)>\mathrm{Hos}\,G$, and let

$$F(G)\geqq L>\mathrm{Hos}\,G\geqq K\geqq 1$$

be a series of normal subgroups of G such that all chief factors between L and K are G-isomorphic to $L/\mathrm{Hos}\,G$ but no chief factor below K is such. Choose K_1 with $L\geqq K_1>K$ where K_1/K is a G-chief factor. By the nilpotency of L and by (1) the chief factor K_1/K can be "permuted" with each one below K. Similarly we can proceed with a chief factor K_2/K_1 below L and so on. Finally we obtain a homogeneous normal subgroup M of G with $M\cong_G L/K$. But then $L=M\times K$ is a product of homogeneous normal subgroups of G which is impossible. Since (1) is hereditary to factor groups, (2) follows.

(2)⇒(1). L/N is nilpotent and therefore contained in $F(G/N)=\mathrm{Hos}\,(G/N)$. But then L/N is a direct product of homogeneous normal subgroups which yields the assertion (1).

Finally we have $N_{\min}(G/N)\leqq F(G/N)$ for each $N\trianglelefteq G$. Hence condition (2) yields $N_{\min}(G/N)\leqq\mathrm{Hos}\,(G/N)$, and this is condition (3) of Theorem 4.6.

4.8. Proposition. *Let G be a solvable group, which has only non-G-isomorphic chief factors in a chief series. Then $\mathfrak{N}(G)\cong\mathrm{Low}\,\mathfrak{H}(G)$ iff in each factor group of G the Fitting group coincides with the socle.*

Proof. Obviously in each factor group of G the homogeneous socle coincides with the socle; further, $\mathfrak{N}_{\mathrm{sep}}(G)=\mathfrak{N}(G)$.

If $F(G/N)=\mathrm{Soc}\,(G/N)$ for each normal subgroup N of G, then in view of Theorems 4.7 and 4.6 we obtain $\mathfrak{N}(G)\cong\mathrm{Low}\,\mathfrak{H}(G)$.

Conversely let $\mathfrak{N}(G)\cong\mathrm{Low}\,\mathfrak{H}(G)$. We take a piece $L>M>N$ of a chief series of G such that L/N is a prime power group. Then $C_G(M/N)\geqq L$, so that no chief factor G-isomorphic to L/M can appear above $C_G(M/N)$. Hence $[L/M]\ngeqq[M/N]$. By Theorem 4.6 L/N decomposes completely as a G-group. This implies in view of Theorem 4.7 the coincidence of the Fitting group and the socle in each factor group of G.

Remark. Recently P. P. Pálfy has shown in [4] that the property (2) in Lemma 1.9 completely describes the prime-coloured poset of a solvable group; i.e. each finite poset, endowed with primes such that (2) of Lemma 1.9 holds, can be represented as the coloured poset $\mathfrak{H}(G)$ of a suitable finite solvable group G.

The author is very grateful to the referee for some useful suggestions.

References

[1] P. Hall, G. Higman, On the p-length of p-soluble groups and reduction theorems for Burnside's problem, *Proc. London Math. Soc.*, **6** (1956), 1—40.
[2] B. Huppert, *Endliche Gruppen. I*, Springer (Berlin—Heidelberg—New York, 1967).
[3] R. Kochendörffer, Über treue irreduzible Darstellungen endlicher Gruppen, *Math. Nachr.*, **1** (1948), 25—39.
[4] P. P. Pálfy, On partial ordering of chief factors in solvable groups, *Manuscripta Math.*, **55** (1986), 219—232.
[5] G. Pazderski, Über lineare auflösbare Gruppen, *Math. Nachr.*, **45** (1970), 1—68.
[6] G. Pazderski, The orders to which only belong metabelian groups, *Math. Nachr.*, **95** (1980), 7—16.
[7] G. Pazderski, Solvable abnilpotent groups, *Math. Nachr.*, **117** (1984), 305—321.

SEKTION MATHEMATIK
WILHELM-PIECK-UNIVERSITÄT ROSTOCK
UNIVERSITÄTSPLATZ 1
2500 ROSTOCK, DDR

Acta Sci. Math., 51 (1987), 185—189

Выпуклые комбинации бесконечных матриц отображений

Л. И. ПОЛОЦКИЙ, М. В. САПИР, Л. А. СКОРНЯКОВ

Посвящается памяти А. П. Хуна

В работе обобщается на бесконечный случай известная теорема Биркгофа о представлении дважды стохастических матриц в виде выпуклой комбинации матриц перестановок ([1], с. 58, теорема 4.8). Этот результат можно рассматривать как еще одно решение 111 проблемы Биркгофа (ср. [3]—[6]). На бесконечный случай обобщается и теорема о представлении стохастической матрицы в виде выпуклой комбинации матриц отображений ([2], с. 493, лемма 5).

Если X и Y — некоторые непустые множества, то $(X\times Y)$-матрицей называется любое отображение A прямого произведения $X\times Y$ в множество действительных чисел. Под X-матрицей понимается $(X\times X)$-матрица. Для матрицы A через A^* обозначается транспонированная матрица, то есть $(Y\times X)$-матрица, определяемая условием $A^*(y,x)=A(x,y)$ для любых x из X и y из Y. Назовем $(X\times Y)$-матрицу квазистохастической матрицей веса α, если выполнены следующие условия:

$$(+)\quad A(x,y)\geqq 0 \quad \text{для любых } x \text{ из } X \text{ и } y \text{ из } Y;$$

$$(\mathrm{S})\quad \sum_y A(a,y)=\alpha \quad \text{для каждого } a \text{ из } X.$$

Квазистохастическая матрица веса 1 называется стохастической. По любому отображению $\beta\colon X\to Y$ можно построить $(X\times Y)$-матрицу $M(\beta)$, полагая

$$M(\beta)(X,Y)=\begin{cases}1, & \text{если } \beta(x)=y,\\ 0, & \text{если } \beta(x)\neq y.\end{cases}$$

Эта матрица, очевидно, является стохастической. Матрицы вида $M(\beta)$ будем называть матрицами отображений. Квазистохастическая X-матрица A веса

Поступило 22 июня 1986 г.

α называется дважды квазистохастической матрицей веса α, если A^* также является квазистохастической матрицей веса α. Если при этом $\alpha=1$, то A называется дважды стохастической матрицей. Если β — это перестановка множества X, то $M(\beta)$, как нетрудно видеть, дважды стохастическая матрица. Она называется матрицей перестановки. Если $A, B, \ldots, D$ — $(X\times Y)$-матрицы, а $\alpha, \beta, \ldots, \delta$ — положительные действительные числа, то матрица

$$F = \alpha A + \beta B + \ldots + \delta D \tag{1}$$

называется положительной комбинацией матриц $A, B, \ldots, D$. Если при этом $\alpha+\beta+\ldots+\delta=1$, то F называется выпуклой комбинацией этих матриц. Далее, если $A, B, \ldots, D$ — [дважды] квазистохастические $(X\times Y)$-матрицы, то и F будет такой же, а если комбинация выпуклая и матрицы [дважды] стохастические, то F — [дважды] стохастическая матрица. Последнее утверждение, очевидно, можно обратить: если F и $A, B, \ldots, D$ — [дважды] стохастические, то комбинация (1) выпуклая. Содержанием матрицы A назовем множество

$$\{\alpha | A(x, y) = \alpha, \quad \alpha \neq 0 \quad \text{для некоторых } x, y\},$$

то есть множество всех ненулевых чисел, содержащихся в A. Условимся для любых двух $(X\times Y)$-матриц A и B писать $A \leqq B$, если $B(x,y)=0$ влечет $A(x, y)=0$ для любых x, y.

Теорема 1. *[Дважды] стохастическая матрица является выпуклой комбинацией матриц [перестановок] отображений тогда и только тогда, когда ее содержание конечно.*

Теорема 1'. *[Дважды] квазистохастическая матрица является положительной комбинацией матриц [перестановок] отображений тогда и только тогда, когда ее содержание конечно.*

В силу сделанных выше замечаний теоремы 1 и 1' эвивалентны, поэтому мы будем доказывать только теорему 1'. Необходимость ее условия очевидна. Доказательству достаточности предпошлем три леммы.

Лемма 1. *Пусть A — [дважды] квазистохастическая $(X\times Y)$-матрица. Тогда найдется матрица [перестановки] отображения $M(\beta)$, для которой $M(\beta)\leqq A$.*

Доказательство. В случае, когда матрица A квазистохастическая, утверждение леммы очевидно: для любого x из X полагаем $\beta(x)$ равным одному из таких y, что $A(x,y)\neq 0$. Если же A дважды квазистохастическая, то применяем [7], с. 693, теорема 2.

Лемма 2. *Если содрежание [дважды] стохастической матрицы A конечно и состоит из рациональных чисел, то A представима в виде положительной комбинации матриц [перестановок] отображений.*

Доказательство. Матрица A представима в виде положительной комбинации матриц $B, C, \ldots, D$ тогда и только тогда, когда матрица nA представима в таком виде, где n — любое положительное число. Поэтому можно считать, что все числа в матрице A целые. Проведем индукцию по весу α матрицы A. База индукции тривиальна: при $\alpha=1$ матрица A сама является матрицей [перестановки] отображения. В силу леммы 1 найдется матрица [перестановки] отображения $M(\beta)$, такая, что $M(\beta)<A$. Разность $A-M(\beta)$ является [дважды] квазистохастической матрицей веса $\alpha-1$. Следовательно, по предположению индукции $A-M(\beta)$ представима в виде положительной комбинации матриц [перестановок] отображений. Отсюда вытекает, что и матрица A представима в таком виде. Лемма доказана.

Всюду далее множество действительных чисел R рассматривается как пространство над полем рациональных чисел Q. Пусть L подпространство в R, Z — базис в L, а P — подмножество из L. Тогда базис Z назовем допустимым для P, если Z состоит из положительных чисел и каждое число из P выражается через Z с положительными коэффициентами.

Лемма 3. *Пусть L — конечномерное подпространство пространства действительных чисел, P — конечное подмножество положительных чисел из L. Тогда в L найдется базис, допустимый для P.*

Доказательство. Воспользуемся индукцией по числу элементов в P. Если в P один элемент, то можно его дополнить до базиса в L положительными числами. Ясно, что полученный базис допустим для P. Пусть для некоторого множества P найден допустимый базис Z. Возьмем произвольное положительное число p из $L \setminus P$. Надо найти базис в L, допустимый для $P \cup \{p\}$. Имеем:

$$p = \sum_1^m \alpha_i a_i - \sum_1^n \beta_j b_j$$

для некоторых a_i, b_j из Z и положительных рациональных α_i, β_j. Пусть d — наименьший общий знаменатель чисел α_i, β_j $(1 \leqq i \leqq m,\ 1 \leqq j \leqq n)$. Можно считать, что базис Z выбран так, что число $\sum d\beta_j$ наименьшее из возможных. Надо доказать, что это число равно 0 и, следовательно, базис Z допустим для $P \cup \{p\}$. Пусть это не так. Поскольку $p>0$, имеем:

$$\sum d\alpha_i a_i > d\beta_1 b_1 \geqq b_1,$$

откуда

$$\sum d\alpha_i a_i / b_1 > 1.$$

Разумеется, существуют такие рациональные положительные числа π_i $(1 \leqq i \leqq m)$, что $\pi_i < a_i/b_1$ и $k = \sum d\alpha_i \pi_i > 1$. Положим $\delta_i = \pi_i/k$. Тогда $\sum d\alpha_i \delta_i > 1, \delta_i < a_i/b_1$. Заменим каждое число a_i из Z на $a_i' = a_i - \delta_i b_1$. Получим новый базис Z' пространства L. Так как $\delta_i > 0$, то каждое число, выражающееся с положительными рациональными коэффициентами через Z, выражается через Z' также с положительными рациональными коэффициентами. Поэтому базис Z' допустим для P. Имеем далее:

$$p = \sum \alpha_i a_i' - \beta_1' b_1 - \beta_2 b_2 - \ldots - \beta_n b_n, \quad \text{где} \quad \beta_1' = \beta_1 - 1/d.$$

Сумма $d\beta_1' + d\beta_2 + \ldots + d\beta_n$ меньше суммы $\sum d\beta_j$. Поскольку $\beta_1 d \geqq 1$, то $\beta_1' \geqq 0$, а так как наибольший общий делитель d' чисел $\beta_1', \beta_2, \ldots, \beta_n$ не превосходит d, то сумма $d'\beta_1' + \sum_{i=2}^{n} d'\beta_i$ меньше суммы $\sum d\beta_i$, что противоречит выбору базиса Z. Лемма доказана.

Теперь у нас все готово для доказательства теоремы 1'. Возьмем [дважды] квазистохастическую X-матрицу A веса α с конечным содержанием P. Рассмотрим подпространство L, натянутое на множество P. В силу леммы 3 в L существует базис Z, допустимый для P. Обозначим через π_z проекцию L на подпространство, натянутое на элемент z из Z. Тогда $A = \sum_{z \in Z} \pi_z(A)$. Пусть π — одна из проекций π_z, $B = \pi(A)$. Поскольку базис Z допустим для P, все числа в матрице B положительные. Возьмем произвольный x из X. Поскольку

$$\sum_y B(x, y) = \pi\Big(\sum_y A(x, y)\Big) = \pi(\alpha) \quad \text{и} \quad \sum_y B(y, x) = \pi\Big(\sum_y A(y, x)\Big),$$

матрица B является [дважды] квазистохастической. Все элементы матрицы B принадлежат одномерному подпространству, натянутому на некоторый элемент z из Z. Поэтому $B = zB'$, где B' — [дважды] квазистохастическая матрица, содержание которой конечно и состоит из рациональных чисел. В силу леммы 2 матрица B представима в виде положительной комбинации матриц [перестановок] отображений F_{zi}. Следовательно, матрица A также представима в виде положительной комбинации матриц F_{zi} по всем z из Z и всем i, что и требовалось.

Теоремы 1' и 1 доказаны.

Замечание. Из цитированной выше теоремы Биркгофа о представимости любой дважды стохастической матрицы конечных размеров в виде выпуклой комбинации матриц перестановок нельзя извлечь никаких оценок числа слагаемых в этом представлении. Из доказательства теоремы 1' вытекает, что это число зависит от содержания матрицы, но не зависит от ее размеров. Действительно, пусть A — дважды стохастическая матрица с конечным содержанием P, L — подпространство, натянутое на P, а Z — допустимый базис для P. Пусть k — число элементов в P. При доказательстве теоремы 1' показано, что матрица A равна положительной комбинации не более, чем k дважды квазистохастических матриц, в которых все числа рациональны (и зависят только от P). Если же в дважды квазистохастической матрице B все числа рациональны и d — их наименьший общий знаменатель, то, как показывает доказательство леммы 2, число матриц перестановок, необходимых для представления B в виде положительной комбинации, не превосходит максимального числа матрицы dB, которое зависит от P, но не от размеров матрицы A.

В заключение авторы хотели бы поблагодарить У. Э. Кальюлайда, обратившего их внимание на работу [7].

Литература

[1] F.-J. Fritz, B. Huppert, W. Willems, *Stochastische Matrizen*, Springer-Verlag (Berlin—Heidelberg—New York, 1979).

[2] G. Hognas, Random semigroup asts on a finite set, *J. Austral. Math. Soc.*, 23 (1977), 481—498.

[3] J. R. Isbell, Birkhoff's problem 111, *Proc. Amer. Math. Soc.*, 6 (1955), 217—218.

[4] D. G. Kendall, On infinite double stochastic matrices and Birkhoff's problem 111, *J. London Math. Soc.*, 35 (1960), 81—84.

[5] B. A. Rattray, J. L. Peck, Infinite stochastic matrices, *Trans. Roy. Soc. Canada*, Sec. 3, 49 (1955), 55—57.

[6] P. Reverz, A probabilistic solution of problem 111 of G. Birkhoff, *Acta Math. Acad. Sci. Hungar.*, 13 (1962), 187—198.

[7] H. Perfect, H. Mirsky, The distribution of positive elements in doubly-stochastic matrices, *J. London Math. Soc.*, 40 (1965), 689—698.

(М. В. С.)
БОТАНИЧЕСКАЯ 23—31
620 083 СВЕРДЛОВСК, СССР

(Л. А. С.)
МОСКОВСКИЙ УНИВЕРСИТЕТ ИМ. М. В. ЛОМОНОСОВА
МЕХМАТ.
119 899 МОСКВА, СССР

Acta Sci. Math., 51 (1987), 191—207

Effective constructions of cutsets for finite and infinite ordered sets

IVAN RIVAL and NEJIB ZAGUIA

Dedicated to the memory of András P. Huhn

Perhaps the most important result in the theory of ordered sets is the 'chain decomposition theorem' of R. P. DILWORTH [1] which states that *in a finite ordered set the minimum number of maximal chains whose union is the set equals the maximum size of an antichain.* However, this maximal antichain need not meet all of the maximal chains in the ordered set. For instance $\{a, d\}$ is an antichain of maximum size in the ordered set N illustrated in Figure 1, and yet $\{a, d\}$ does not meet the maximal chain $\{c, b\}$.

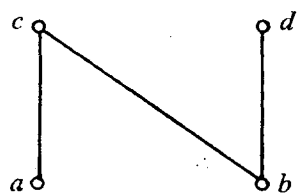


Figure 1

Call a subset K of an ordered set P a *cutset* of P if every maximal chain of P meets K. If K is an antichain then we call it an *antichain cutset* of P. If $K-\{x\}$ is not a cutset for every x in K then we call it a *minimal cutset* of P.

The N illustrated in Figure 1 is the union of the antichain cutsets $\{a, b\}$ and $\{c, d\}$. In contrast the ordered set illustrated in Figure 2 cannot be the union of antichain cutsets at all, since there is no antichain cutset which contains x.

I. RIVAL and N. ZAGUIA [4] have shown that *a finite ordered set is the union of antichain cutsets if and only if it contains no alternating-cover cycle.* For $n \geqq 2$, a

Received April 24, 1986, and in revised form November 10, 1986.

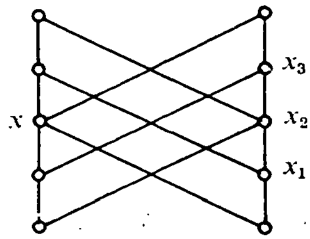


Figure 2

subset $\{x, a_1, c_1, \ldots, a_n, c_n\}$ of an ordered set P is called a *generalized alternating-cover cycle* (for x) provided that

$$c_1 > x > a_n, \quad c_1 > a_1, \quad c_2 > a_2, \ldots, c_n > a_n,$$

$$c_1 > a_n, \quad c_2 > a_1, \ldots, c_{n-1} > a_{n-2}, \quad c_n > a_{n-1}$$

and provided that

$$c_1 \succ a_1, c_2 \succ a_2, \ldots, c_n \succ a_n$$

are covering relations in P itself. If these are the only comparabilities among the elements $\{x, a_1, c_1, \ldots, a_n, c_n\}$, we call this ordered set an *alternating-cover cycle* (see Figure 3). For emphasis we sometimes indicate by 'double lines' in the figures the covering relations in an ordered set. We also say that x is contained in a generalized alternating-cover cycle. Actually I. Rival and N. Zaguia [4] prove this more general result.

Theorem 1. *In a finite ordered set an element is contained in an antichain cutset if and only if it is not contained in a generalized alternating-cover cycle.*

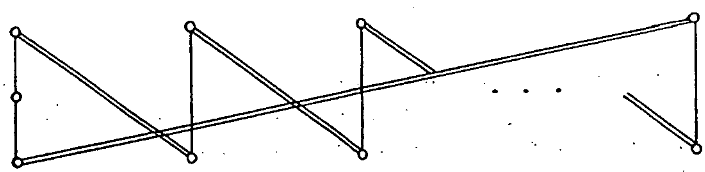


Figure 3

Here, we give an 'efficient' algorithm for the construction of an antichain cutset containing a given element, which is in effect another proof of the same theorem. From the order-theoretical point of view, the first proof in I. Rival and N. Zaguia [4] is certainly shorter and perhaps more elegant. Still, the algorithm for the construction of the antichain cutset implicit in that proof, seems on the surface at least to perform in an 'exponential' number of steps, as a function of the number of elements of the ordered set. Another proof, presented here is algorithmically much better. Indeed, the algorithm for the construction of the antichain cutset, implicit in this proof, performs in a 'polynomial' number of steps as a function of the number of elements of the ordered set.

In the ordered set illustrated in Figure 2, the element x is not contained in any antichain cutset. Still there is a minimal cutset $\{x, x_1, x_2, x_3\}$ which contains x. In a finite ordered set P, every element is contained in a minimal cutset, and thus a finite ordered set P is always the union of minimal cutsets. That is not always the case for infinite ordered sets though. For instance, in $2\times(\omega+1)$ (see Figure 4), there is no minimal cutset which contains x even though, for example every chain in $2\times(\omega+1)$ has a supremum and an infimum.

An ordered set P is *regular* if every nonempty chain C of P has a supremum and an infimum and, whenever $x<\sup C$ (respectively, $x>\inf C$), $x<c$ (respectively $x>c$), for some element c in C. We expect that regular ordered sets can be expressed as the union of minimal cutsets but we are unable to prove that yet. Here is a partial solution.

Theorem 2. *A regular ordered set satisfying a chain condition is the union of minimal cutsets.*

An ordered set is said to satisfy a *chain condition* if it does not contain either an infinite, strictly descending chain $x_1>x_2>\dots$ or it does not contain an infinite, strictly ascending chain $x_1<x_2<\dots$.

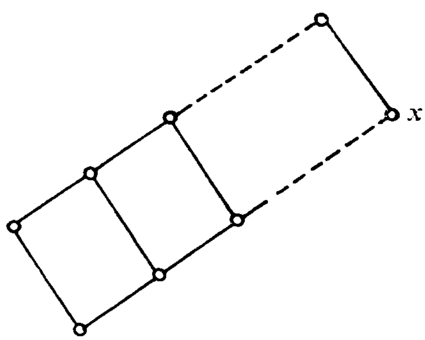


Figure 4

In general, though, a minimal cutset need not be an antichain and, of course, an antichain need not be a (minimal) cutset. I. Rival and N. Zaguia [4] have shown that, in an ordered set which contains no subset isomorphic to N, every finite, minimal cutset is an antichain. D. Higgs [3] has extended this result to arbitrary minimal cutsets, and has also proved the converse in the case of finite ordered sets. As a consequence of Theorem 2, we can extend Higgs's result.

Theorem 3. *Let P be a regular ordered set satisfying a chain condition. Then, every minimal cutset in P is an antichain if and only if P contains no subset isomorphic to N.*

D. Higgs [3] was the first to give an example of an ordered set which contains subsets isomorphic to N and in which every minimal cutset is an antichain.

A related question is whether every maximal antichain meets every maximal chain, that is, whether every maximal antichain is a cutset? The ordered set illustrated in Figure 5, has a maximal antichain which is not a cutset.

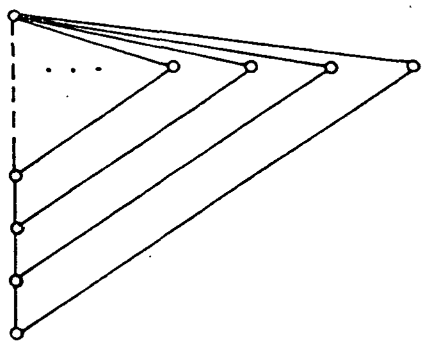


Figure 5

An early result of P. A. Grillet [2] shows that, *for a regular ordered set P, every maximal antichain meets every maximal chain if and only if P contains no subdiagram isomorphic to N.* We extend this result in terms of 'generalized' N's. This gives a characterization of those ordered sets in which every maximal antichain is a cutset.

Let A_1 and A_2 be subsets of an ordered set P. Write $A_1 < A_2$ if, for every $u \in A_1$ and $v \in A_2$, $u < v$. We say that A_2 *covers* A_1 (or A_1 is *covered* by A_2) and write $A_1 \prec A_2$ if $A_1 < A_2$ and there is no x in P such that $A_1 < \{x\} < A_2$. Also, we use the convention that $\emptyset < A$ for every subset A of P. We say that A_1 is *cofinal* for A_2 (respectively *coinitial*) provided that, for every $v \in A_2$, there exists $u \in A_1$ such that $v \leqq u$ (respectively $u \leqq v$). Let C_1, C_2, A_1 and A_2 be subsets of P such that C_1 and C_2 are chains in P and $A_1 \cup A_2$ is an antichain in P. We call the four-tuple (C_1, C_2, A_1, A_2) a *generalized N* provided that $C_1 \prec C_2$ and A_1 is cofinal for C_1 and A_2 is coinitial for C_2. In Figure 6, we illustrate basic examples of generalized N's. An N, too, is a generalized N. Also, in a regular ordered set, it is easy to see that every generalized N must be an N in the diagram itself.

Theorem 4. *In an ordered set every maximal antichain meets every maximal chain if and only if it contains no generalized N as a subdiagram.*

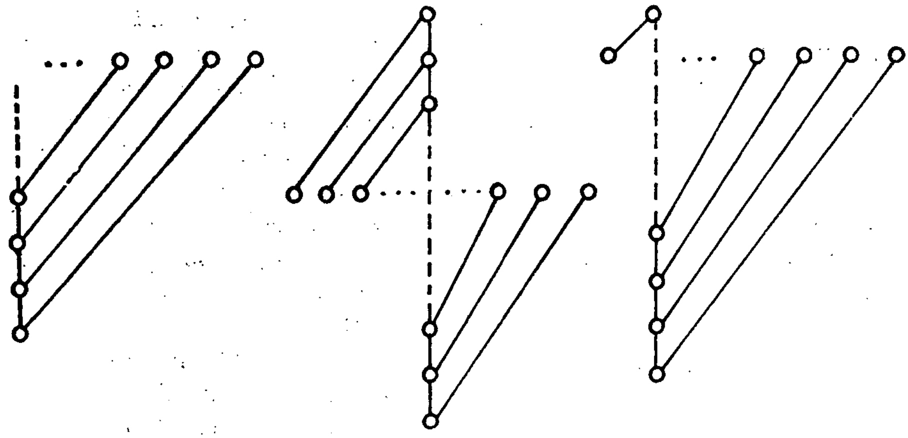


Figure 6

Proof of Theorem 1. For purposes of the proof it is convenient to speak of cutsets "for" elements. Say that a subset S is a cutset for x if $S\cup\{x\}$ is a cutset and each $s\in S$ is noncomparable with x. We shall prove, by induction on the cardinality of P, that every element has an antichain cutset provided that no element of P is contained in a generalized alternating-cover cycle. To this end let x be an element of P with no antichain cutset in P. Then x cannot be a maximal element of P for then we could choose $A(x)$ to be the set of maximal elements of P distinct from x. Let u be a maximal of P, $u>x$, and put $P'=P-\{u\}$. Notice that no element of P' is contained in a generalized alternating-cover cycle so, by the induction hypothesis, x must have an antichain cutset $A'(x)$ in P'. We may suppose that $A'(x)$ is not an antichain cutset for x in P. Then there is a maximal chain $C(u)$ of P which contains u but no element from $A'(x)\cup\{x\}$. Let u' be the lower cover of u in $C(u)$. According to the induction hypothesis any maximal chain in P' containing $C(u)-\{u\}$ must contain some element from $A'(x)\cup\{x\}$. Therefore, $u'\leqq x$ or $u'\leqq v$ for some v belonging to $A'(x)$. But $u'\not\leqq x$ since $u>x$ and $u'\prec u$; therefore, $u'\leqq v$ for some v in $A'(x)$. Our aim is to construct, starting from $A'(x)$, an antichain cutset for x in P. We cannot hope to use u in an antichain cutset for x. In order to "account" for the maximal chain $C(u)$ we may, however, try to use u', but then we could not use v in an antichain cutset for x. Then we may seek to replace v by other elements, each noncomparable to x and to the members of the "current" cutset for x.

We shall introduce and develop a "two-player game" which we use later to effect the construction of an antichain cutset. Meet our players: CHAIN — the villain— and, ANTICHAIN— our hero—.

The setting for our spectacle is a finite ordered set P which contains no alternating-cover cycle. Let x belong to P and let $A(x)$ be a minimal cutset for x (that is, a cutset for x such that, for each y in $A(x)$, $A(x)-\{y\}$ is not a cutset for x in P). Notice that for each y in $A(x)$ there is a maximal chain $C(y)$ in P such that $C(y)\cap A(x)=\{y\}$. Call such a maximal chain in P *essential* for y in $A(x)$.

A *down game* between CHAIN and ANTICHAIN (in P for x) is played as follows. CHAIN is the first to move: CHAIN selects an element c_1 from $A_1=A(x)$, for which there is another element a_0 in A_1 such that $c_1>a_0$, if one exists. (In effect, CHAIN "uncovers" evidence why A_1 is not an antichain cutset for x.) If CHAIN has no such move then we say that ANTICHAIN *wins the down game* (indeed, this must mean that A_1 is an antichain cutset for x). Otherwise ANTICHAIN responds in this down game by identifying all lower covers of c_1 on essential chains for c_1 in A_1, each of which is not below x itself: call these elements $a_1^1, a_1^2, \ldots$. These elements constitute ANTICHAIN's first move in reply to CHAIN's move c_1. We now "reform" the cutset A_1 by constructing a minimal cutset A_2 for x contained in

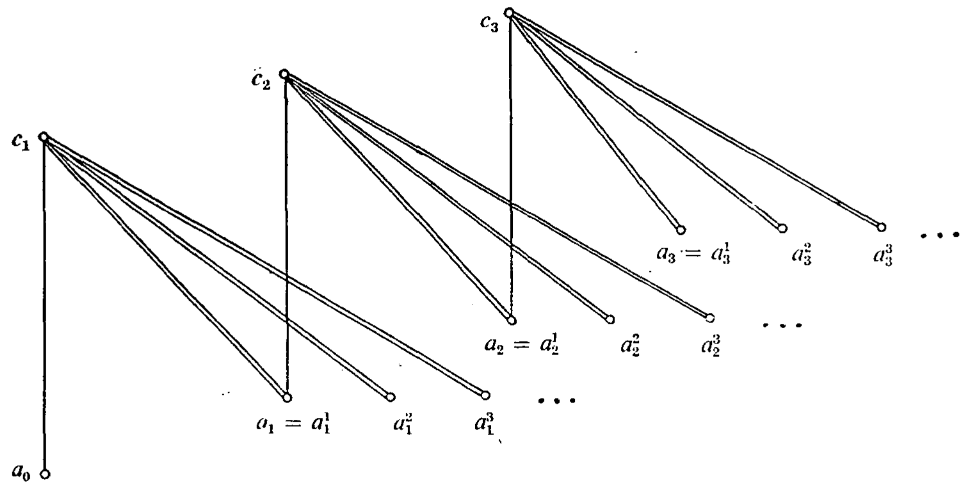
$$\{a_1^1, a_1^2, \ldots\}\cup(A_1-\{c_1\}).$$

Evidently, A_2 consists of two disjoint subsets: the first consists of the sequence $a_1^1, a_1^2, \ldots$ which is an antichain; the second is a subet of $A_1-\{c_1\}$, which together with the sequence of elements constituting ANTICHAIN's move is a minimal cutset for x. Notice that just as CHAIN may not be able to move (if A_1 is already an antichain cutset), it may be that ANTICHAIN cannot respond to CHAIN's move c_1: this would be the case if some lower cover of c_1 on an essential chain for c_1 in A_1 is itself below x, such an element cannot be in a cutset for x. If, then, ANTICHAIN cannot move we say that CHAIN *wins the down game.* If both first moves can be made then the down game continues. CHAIN selects an element c_2 from A_2 such that $c_2>a_1$, where a_1 belongs to $\{a_1^1, a_1^2, \ldots\}$. If CHAIN has no such move then ANTICHAIN wins the down game. Otherwise, ANTICHAIN again responds by selecting all lower covers of c_2 on essential chains for c_2 in A_2, each of which is not below x, say $a_2^1, a_2^2, \ldots$ CHAIN's move c_2. Again the cutset A_2 is altered to construct a minimal cutset A_3 for x contained in

$$\{a_2^1, a_2^2, \ldots\}\cup(A_2-\{c_2\})$$

(see Figure 7). Again the minimal cutset contains the antichain consisting of the elements in ANTICHAIN's second move and, as well, a subset of $A_2-\{c_2\}$. If CHAIN can now move then CHAIN will choose an element c_3 from A_3 such that $c_3>a_2$, where a_2 belongs to $\{a_2^1, a_2^2, \ldots\}$. And so on.

Furthermore, by a sequence $G_1, G_2, \ldots$ of down games we mean that each of the down games G_i begins with a comparability taken from the current cutset at the end of the preceding down game G_{i-1}, for each $i=2, 3, \ldots$.



Construction in the proof of Theorem 1.
Figure 7

We call the cutset A_k, $k=1, 2, \ldots$, the current cutset for x at CHAIN's k^{th} move, and then ANTICHAIN's k^{th} move. We say that ANTICHAIN *wins this down game* if, for some $k \leqq |P|$, CHAIN cannot make a k^{th} move in this down game; otherwise, we say that CHAIN *wins this down game.*

An up game between CHAIN and ANTICHAIN is defined dually. In an up game CHAIN's k^{th} move is to select an element c_k' from the current cutset A_k' such that $c_k' < a_{k-1}'$, where a_{k-1}' is one of the elements

$$a_{k-1}'^{1}, a_{k-1}'^{2}, \ldots$$

chosen by ANTICHAIN in the $k-1^{\text{th}}$ move.

Let G be a down game for x in P and let $c_1 > a_0$ be the first move for CHAIN. We say that the down game G is linked to x provided that there are sequences

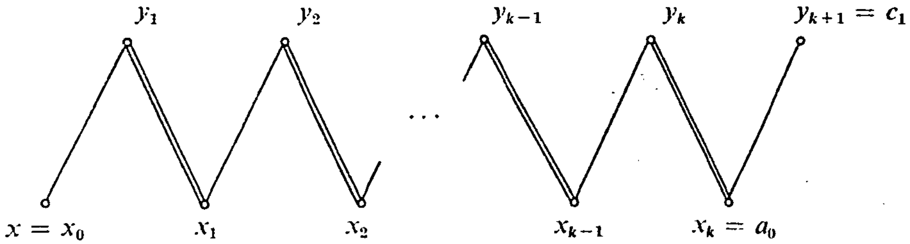
$$x = x_0, x_1, x_2, \ldots, x_k = a_0, \quad y_1, y_2, \ldots, y_k$$

in P such that

$$x_i < y_{i+1} \quad \text{and} \quad x_{i+1} \prec y_{i+1}$$

for each $i=0, 1, 2, \ldots, k-1$ (see Figure 8). The notion of an up game linked to x is defined dually.

We shall now establish two technical lemmas needed for the proof of the Theorem. The first shows that once CHAIN moves in a down game, that move can never be repeated in that down game. Moreover, in any sequence of down games, once CHAIN moves in one of the down games, that move can never be repeated in any



Construction in the proof of Theorem 1.
Figure 8

later move of any later down game. To this end we write $(m, j) < (n, k)$ for integers m, n, j, k provided that, either $m < n$, or else, $m = n$ and $j < k$ (the usual lexicographic order).

Lemma 1. *Let P be a finite ordered set and let x belong to P. Let $G_1, G_2, \ldots$ be a sequence of down games in P for x. Let $A_1^m, A_2^m, \ldots$ be the current cutsets for x in the game G_m. Then*

$$c_j^m \notin A_k^n \quad \textit{whenever} \quad (m, j) < (n, k).$$

Proof of Lemma 1. According to the rules of play, c_j^m does not belong to A_{j+1}^m. Suppose, however, that there are integers m, n, j, k such that $c_j^m \in A_k^n$. Suppose that (k, n) is chosen least with this property in the lexicographic order. This means that c_j^m is an element of the $k-1^{\text{th}}$ move of ANTICHAIN, that is, $c_j^m \prec c_{k-1}^n$. Let C_{k-1}^n denote the essential chain (containing c_j^m) for c_{k-1}^n in A_{k-1}^n and similarly, let C_j^m denote the essential chain (containing a_j^m) for c_j^m in A_j^m. We use these maximal chains C_{k-1}^n and C_j^m to construct another maximal chain C defined by

$$C = (C_{k-1}^n \cap (c_j^m]) \cup (C_j^m \cap [c_j^m))$$

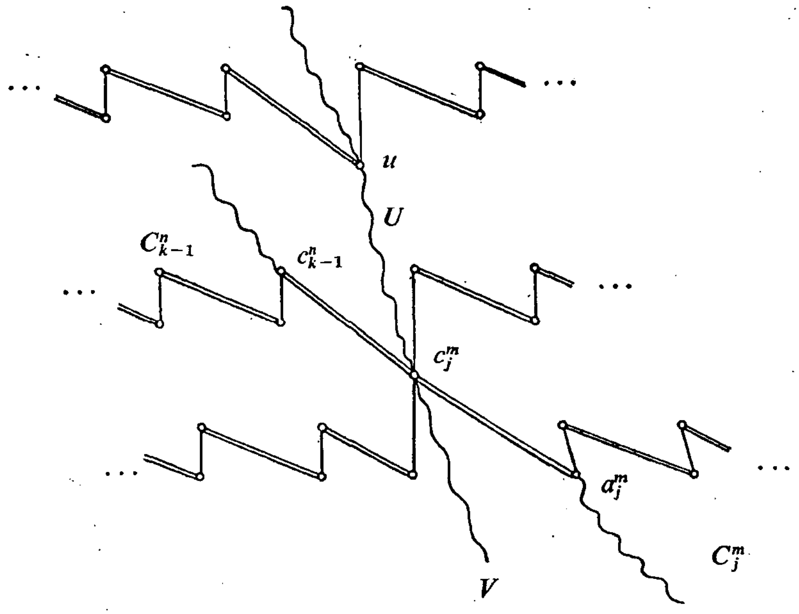
(see Figure 9). (For y in P, $(y] = \{x \in P \mid x \leqq y\}$ and $[y) = \{x \in P \mid x \geqq y\}$.) Let

$$U = C \cap [c_j^m) \quad \text{and} \quad V = C \cap (c_j^m].$$

Suppose the chain U contains an element of A_{k-1}^n. Let $m \leqq p \leqq n$ be the least integer for which there is some i such that A_i^p contains an element of U. Let i, too, be the least integer with this property and choose u a maximal element belonging to $U \cap A_i^p$. Note that u does not belong to A_j^m. Let C_{i-1}^p be the essential chain (containing u) for c_{i-1}^p in A_{i-1}^p. Now construct

$$C' = (C_{i-1}^p \cap (u]) \cup (C_j^m \cap [u)).$$

Evidently A_{i-1}^p cannot contain any element of $C_{i-1}^p \cap (u]$ and, by the maximality of u it cannot contain any element of $C_j^m \cap [u)$, either. That is impossible. It follows that U cannot contain any member of A_{k-1}^n.



Construction in the proof of Lemma 1.
Figure 9

On the other hand, V cannot contain any element of A_{k-1}^n, either. This, in turn, implies that C contains no element of A_{k-1}^n, which is a contradiction.

The next lemma indicates just how the play of games between CHAIN and ANTICHAIN is affected by generalized alternating-cover cycles. Absence of generalized alternating-cover cycles gives ANTICHAIN decided advantage.

Lemma 2. *Let P be a finite ordered set and let x be an element of P which is contained in no generalized alternating-cover cycle. Then* ANTICHAIN *wins every down game (in P for x) linked to x.*

Proof of Lemma 2. Let G be a down game. The first move for CHAIN is an element c_1 for which there is an element a_0 in the current cutset A_1 for x. Suppose that CHAIN wins some down game G. Then, according to Lemma 1, there are (finitely many) distinct elements

$$c_1, c_2, \ldots, c_j$$

(the sequence of CHAIN's moves) and there are elements

$$a_1, a_2, \ldots, a_{j-1}$$

(the sequence of ANTICHAIN's moves) such that CHAIN wins this down game in j moves. Therefore, there must be a lower cover a_j on an essential chain for c_j in A_j such that $a_j < x$. Let

$$x = x_0, x_1, x_2, \ldots, x_k = a_0, \quad y_1, y_2, \ldots, y_{k+1} = c_1$$

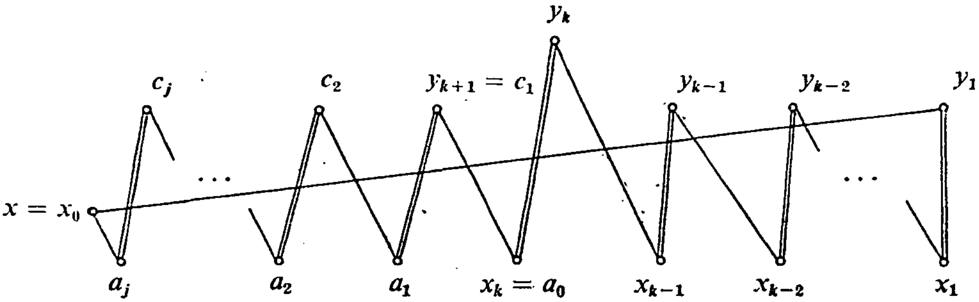
be elements satisfying

$$x_i < y_{i+1} \quad \text{and} \quad x_{i+1} \prec y_{i+1}, \quad i = 0, 1, \ldots, k-1.$$

Then x is contained in the generalized alternating-cover cycle

$$x = x_0, x_1, y_1, x_2, y_2, \ldots, x_k, y_k, a_1, c_1 = y_{k+1}, a_2, c_2, \ldots, a_j, c_j$$

a contradiction (see Figure 10).



Construction in the proof of Lemma 2.
Figure 10

We are now ready to proceed directly with the proof of Theorem 1.

Let P be a finite ordered set and let x be an element of P contained in no generalized alternating-cover cycles. We shall show by induction on the cardinality of P, that x has an antichain cutset. If, for instance, x is a maximal element of P then the remaining maximal elements of P distinct from x constitute an antichain cutset (possibly empty) for x. Let us suppose, then, that x is not a maximal element of P. Let u be a maximal element of P satisfying $u > x$ and put $P' = P - \{u\}$. Of course, P', too, does not contain any generalized alternating-cover cycles for x, so we may apply the induction hypothesis to P' to obtain an antichain cutset A' for x in P. We may suppose that A' is not an antichain cutset for x in P. Then there are maximal chains in P, each containing u and each disjoint from A'. Let u' be the unique lower cover of u on some such maximal chain C. Now $C - \{u\}$ is not a maximal chain in P', for otherwise $C - \{u\}$, whence C itself, contains an element of A'. Then, for every such u' there is some v in A' satisfying $u' < v$.

Since $u > x$, no lower cover of u can lie below x. In fact

$$A' \cup \{u' \mid u' < v \text{ for some } v \text{ in } A'\}$$

is a cutset for x in P. Let A be a minimal cutset for x in P contained in this set. Notice that A contains each of these elements u'. We may suppose that A is not an antichain.

Our aim now is to successively construct new cutsets, at each stage eliminating comparabilities of the form $u'<z$ where $u'\prec u$, $u'<v$ for some v in A' and z belongs to the current cutset. As there are only finitely many such comparabilities and there can be no repetitions this construction must terminate with an antichain cutset.

Here is the first step in the construction. We begin with the comparabilities $u_1'<z_1$ in the current cutset $A=A^{1,0}$. Put $c_1^{1,1}=z_1$ and $a_0^{1,1}=u_1'$. Since $x<u$ and $u\succ u_1'$, $u_1'<z_1$, then according to Lemma 2, ANTICHAIN wins a down game G_1^1 which begins with $a_0^{1,1}<c_1^{1,1}$. Let $A^{1,1}$ be the current cutset at the end of such a down game. Let $a_0^{1,2}\in(A^{1,1}-A^{1,0})$ and $c_1^{1,2}\in A^{1,1}$ satisfy $a_0^{1,2}<c_1^{1,2}$ and let G_2^1 be a down game which begins with $a_0^{1,2}<c_1^{1,2}$. As there are sequences

$$x,\ u_1' = a_0^{1,1}, a_1^{1,1}, a_2^{1,1}, \ldots, a_0^{1,2}$$

and

$$u,\ z_1 = c_1^{1,1}, c_2^{1,1}, \ldots, c_1^{1,2}$$

G_2^1 is linked to x and so ANTICHAIN wins G_2^1. In general, let $a_0^{1,i}\in A^{1,i}-A^{1,0}$ and $c_1^{1,i}\in A^{1,i}$ satisfy $a_0^{1,i}<c_1^{1,i}$ where $A^{1,i}$ is the current cutset at the end of the down game G_{i-1}^1, and let G_i^1 be a down game which begins with $a_0^{1,i}<c_1^{1,i}$. Again, G_i^1 is linked to x, so ANTICHAIN wins G_i^1. By Lemma 1, this sequence terminates after $m(1)$ such successive games. Let $A^{1,m(1)}$ be the current cutset at the end of this sequence of down games.

We may suppose that $A^{1,m(1)}$ is not an antichain. We show that any comparability $y<z$ in $A^{1,m(1)}$ satisfies $y=u'$ for some $u'\prec u$ and $u'<v$ for some $v\in A'$. By Lemma 1, $z\neq z_1$. If $y\notin A^{1,0}$ then this sequence

$$G_1^1, G_2^1, \ldots, G_{m(1)}^1$$

can be extended by a down game $G_{m(1)+1}^1$ which begins with $a_0^{1,m(1)+1}=y<z=c_1^{1,m(1)+1}$. We may suppose therefore that $y\in A^{1,0}\cap A^{1,m(1)}$. Next observe that each element in every move for ANTICHAIN is below some element in $A=A^{1,0}$. To see this we proceed by induction. Evidently, $a_0^{1,0}=u_1'<z_1$ and $z_1\in A=A^{1,0}$. In general, let

$$a_k^{1,i} \prec c_k^{1,i}.$$

If $c_k^{1,i}\in A^{1,0}$ then we are done. Otherwise, $c_k^{1,i}=a_1^{1,j}$ for $(j, 1)<(i, k)$ in the lexicographic order. By induction $a_1^{1,j}$ is below some element of $A^{1,0}$ and so, $a_k^{1,i}$, too, is below some element of $A^{1,0}$. This means that, in particular, there is $t\in A^{1,0}$ satisfying $z\leqq t$ and so $y<t$ in $A^{1,0}$. That in turn, implies that y is a lower cover of u and so y is some u', where $u'<v$ for some $v\in A^{1,0}$.

Let us suppose that we have completed $k-1$ steps in this construction. Let $A^{k,0}=A^{k-1,m(k-1)}$ be the current cutset, let $u'_k<z_k$ be a comparability in $A^{k,0}$ and let G_1^k be a down game beginning with

$$a_0^{k,1} = u'_k < z_k = c_1^{k,1}.$$

Now $x<u$, $u\succ u'_k$ and so G_1^k is linked to x whence ANTICHAIN wins this down game. Let $A^{k,1}$ be the current cutset at the end of G_1^k. Let

$$a_0^{k,2}\in(A^{k,1}-A^{k,0}), \quad c_1^{k,2}\in A^{k,1}$$

satisfy $a_0^{k,2}<c_1^{k,2}$. Again ANTICHAIN wins any down game G_2^k beginning with $a_0^{k,2}<c_1^{k,2}$.

In general, let

$$a_0^{k,i}\in(A^{k,i-1}-A^{k,0}), \quad c_1^{k,i}\in A^{k,i-1}$$

satisfy $a_0^{k,i}<c_1^{k,i}$ where $A^{k,i}$ is the current cutset at the end of the down game G_{i-1}^k. As before the down game G_i^k beginning with $a_0^{k,i}<c_1^{k,i}$ is linked to x, so ANTICHAIN wins this down game. By Lemma 1, this sequence

$$G_1^k, G_2^k, \ldots, G_i^k, \ldots$$

terminates after finitely many such successive down games, $m(k)$ say.

We may suppose that $A^{k,m(k)}$, the current cutset at the end of the down game $G_{m(k)}^k$, is not an antichain. Let $y<z$ in $A^{k,m(k)}$. As the sequence

$$G_1^k, G_2^k, \ldots, G_{m(k)}^k$$

cannot be extended, $y\in(A^{k,0}\cap A^{k,m(k)})$. Again as above, each element in every move for ANTICHAIN in every game G_i^k, is below some element in $A^{k,0}$. Therefore, there is $t\in A^{k,0}$ satisfying $z\leqq t$, so $y<t$ in $A^{k,0}$. By induction, y must be some u', where $u'\prec u$ and $u'<v$ for some v in $A^{1,0}$.

By Lemma 1, there can be no repetition of the comparabilities $y<z$ where y is of the form u' with $u'\prec u$ and $u'<v$ for some v in $A^{1,0}$. As there are only finitely many comparabilities of this type the process must end and the current cutset at the end of this construction must be an antichain. This completes the proof.

Implicit in this proof of Theorem 1 is an effective procedure to construct an antichain cutset. We do this, as in the proof, by a sequence of 'moves'. Every move begins with a comparability in P. According to Lemma 1, two different moves always begin with different comparabilities; thus, at most n^2 moves are needed to produce an antichain cutset. It remains, therefore, to prove that a move can be effected in a polynomial (in n) number of steps. This is the outline of a move.

(i) Find a comparability $a<b$ in the current cutset.

(ii) Replace b in the current cutset by all of its lower covers.

(iii) Remove, from among these lower covers those, nonessential to the new cutset.

(iv) Remove any further elements nonessential to the new cutset.

(v) The new minimal cutset is the current cutset for the next move.

The only outstanding item is how to decide effectively whether or not an element is essential, in a cutset.

Let K be a cutset of an ordered set P and let $x \in K$. Then x is *essential* in K if there is a maximal chain C in P such that $C \cap K = \{x\}$. Let $I(x) = \{y \in P \mid \text{either } y < x \text{ or } y > x\}$. Then x is essential in K if and only if $K \cap I(x)$ is not a cutset in $I(x)$. Our problem therefore reduces (is polynomially equivalent to) the following.

Given a subset K in P is there an effective procedure to decide whether or not K is a cutset of P?

If the subset K is an antichain, then, as is well known, K is a cutset in P if and only if P does not contain an $N = \{a < c,\ b < d,\ b \prec c\}$, such that $\{a, d\} \subseteq K$. Obviously this can be decided in a polynomial number of steps too.

If K is not an antichain, we can transform (polynomially) P to an ordered set P' and K to an antichain K' of P' such that K is a cutset of P if and only if K' is a cutset of P'. To see this we consider several cases. Let $x, y \in K$ satisfy $x \prec y$ in K. If $x \prec y$ in P too then we may delete the covering edge $x \prec y$ to obtain P' and K'. Let us suppose that $x \prec z < y$ in P for some z and suppose that there is no $t < z$ with $t \nleqq x$. In this case we construct P' by only removing the element z and we choose K' in P' to be the same set as K. If, for each $x \prec z < y$, there is $t < z$ with $t \nleqq x$, then we remove the edge $x \prec z$ again to produce the ordered set P', and K' is the same set as K.

Proof of Theorem 2. The proof consists in showing that every element in P is in some minimal cutset of P. Let $x \in P$ and let C_x be a maximal chain of P such that $x \in C_x$.

We suppose that the ordered set P has no infinite decreasing chains. (Otherwise, if P has no infinite increasing chains then we apply dual arguments.) Let

$$B_1 = \{y \in P - C_x \mid y \succ z \text{ for some } z \text{ in } C_x \text{ and } z < x\}$$

and

$$B_2 = \{y \in P - C_x \mid y \text{ is minimal in } P\}.$$

The subset $A = B_1 \cup B_2 \cup \{x\}$ of P is a cutset. Indeed, let C be a maximal chain in P. If $C \cap B_2 = \emptyset$ then $\inf C_x \in C$. Therefore $C_x \cap C \neq \emptyset$. We set

$$u = \sup \{t \in C_x \cap C \mid t < x\}.$$

Since P is chain complete, $u \in C_x \cap C$. If $u = x$ then $x \in C$, otherwise there is v an upper cover of u such that $v \in C - C_x$. Thus $v \in C \cap B_1$. Also, x is essential in A

since $A\cap C_x=\{x\}$. The ordered set P contains no infinite descending chains, so we can consider a well ordering of

$$B_1\cup B_2 = \{x_1, x_2, \ldots, x_\alpha, \ldots\}_{\alpha<\lambda}$$

which is an extension of the order on $B_1\cup B_2$. Thus $x_i<x_j$ in P implies $i<j$.

Now, we define an algorithm which transforms the cutset A to a minimal cutset of P containing x. Let $\{A_1, A_2, \ldots, A_\alpha, \ldots\}_{\alpha\leqq\lambda}$ be a sequence of cutsets of P, defined inductively as follows. If x_1 is essential in A then $A=A_1$. Otherwise $A_1=A-\{x_1\}$. Assume we have already defined $(A_i)_{i<\alpha}$. If the ordinal α is isolated $\alpha=\beta+1$, and x_α is essential in A_β then $A_\alpha=A_\beta$. Otherwise $A_\alpha=A_\beta-\{x_\alpha\}$. If the ordinal α is a limit $\alpha=\sup_{\beta<\alpha}\beta$, then $A_\alpha=\bigcap_{\beta<\alpha} A_\beta$.

First of all, we prove by induction on α that A_α is a cutset of P for every $\alpha<\lambda$. Suppose that α is the least ordinal such that A_α is not a cutset of P. Let C be a maximal chain in P such that $C\cap A_\alpha=\emptyset$. If $\alpha=\beta+1$, then A_β is a cutset of P, thus $A_\beta\cap C\neq\emptyset$. From the inductive construction of A_α, $A_\alpha-A_\beta\subseteqq\{x_\alpha\}$, therefore $A_\beta\cap C=\{x_\alpha\}$ which means that x_α is essential in A_β. So $x_\alpha\in A_\alpha$, which is a contradiction. If $\alpha=\sup_{\beta<\alpha}\beta$, then $A_\alpha=\bigcap_{\beta<\alpha} A_\beta$ and $C\cap A_\beta\neq\emptyset$ for every $\beta<\alpha$. (Thus $|A\cap C|$ is infinite, for otherwise, let x_μ be in $A\cap C$ with a largest index μ, thus $x_\mu\notin A_\mu$ which gives $A_\mu\cap C=\emptyset$.) Also $A\cap C$ contains infinitely many elements in B_1, since C cannot contain more than one element in B_2 (B_2 is an antichain). Let

$$C\cap B_1 = \{y_1, y_2, \ldots, y_i, \ldots\}.$$

Then for every i, y_i covers t_i for some t_i in C_x and $t_i<x$. Let $y=\sup y_i$ and $t=\sup t_i$. Since $y>t_i$ for every i and P is chain complete, $y\geqq t$. If $y>t$, then from the regularity of P, $t<y_j$ for some j. Therefore $t_j<t<y_j$ which contradicts $t_j\prec y_j$. Thus $y=t$ and $y\in C_x\cap C$. Obviously $y\neq x$ since $C\cap A_\alpha=\emptyset$ and $x\in A_\alpha$. Also if $y>x$, then from the regularity of P, $x<y_j$ for some j, thus $t_j<x<y_j$, which contradicts $y_j\succ t_j$. Therefore $y<x$. Now consider the maximal chain

$$K=(C_x\cap(y])\cup(C\cap[y))$$

of P. Obviously $K\cap A=\emptyset$, which contradicts that A is a cutset of P.

The subset A_λ is a minimal cutset. Indeed let $x_\alpha\in A_\lambda$ then $x_\alpha\in A_\beta$, for every $\beta<\lambda$. In particular $x_\alpha\in A_\alpha$, which implies the existence of a maximal chain C of P such that $C\cap A_\alpha=\{x\}$. Since $A_\lambda\subseteqq A_\alpha$, $C\cap A_\lambda=\{x_\alpha\}$. This completes the proof.

The proof of Theorem 2, does not extend to the case that P does not satisfy a chain condition. Indeed, in the example illustrated in Figure 11, $\{x\}\cup B_1\cup B_2$ is not a cutset.

In general, it need not be the case that a cutset always contains a minimal one, even for regular ordered sets. For instance, the ordered set illustrated in Figure 12

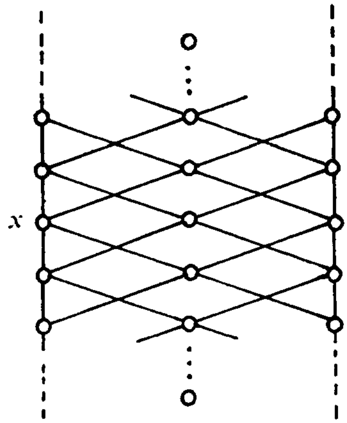


Figure 11

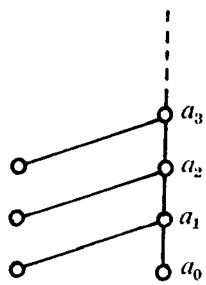


Figure 12

is regular and $\{a_1, a_2, \ldots\}$ is a cutset which does not contain a minimal one. A related question is this. Which ordered sets contain at least one minimal cutset?

Proof of Theorem 3. Suppose that P contains a subset $\{a, b, c, d\}$ isomorphic to N, that is, $a<c$, $b<c$, $b<d$ are the only comparabilities among the elements $\{a, b, c, d\}$. Without loss of generality, we may assume that $c \succ a$ and $d \succ b$. Consider maximal chains C and D of P such that $\{a, c\} \subseteqq C$ and $\{b, d\} \subseteqq D$. Obviously $C \cap \{b, d\} = \emptyset$ and $D \cap \{a, c\} = \emptyset$. Assume that for every x in D and $x \geqq d$, x does not cover in P any element y such that $y \in C$ and $y \leqq a$. Thus $\{c, b\}$ is a cutset in $P_0 = C \cup D$.

From the proof of Theorem 2, there is a minimal cutset K of P containing c such that $K \subseteqq \{c\} \cup B_1 \cup B_2$ with

$$B_1 = \{y \in P - C \mid y \succ z \text{ for some } z \text{ in } C \text{ and } z < c\}$$

and

$$B_2 = \{y \in P - C \mid y \text{ is minimal in } P\}.$$

Since K is a cutset, $K\cap D\neq\emptyset$. Let $x\in K\cap D$. If $x\in B_1$, then $x\succ z$ in P for some z in C and $z<c$. By assumption $x\leqq b$, so $x<c$. If $x\in B_2$, then $x=\inf D$ thus $x<c$. Therefore K is not an antichain.

Assume that there exists x in D with $x\geqq d$ and such that x covers y in P for some $y\in C$ and $y\leqq a$. Necessarily, either $x\neq d$ or $y\neq a$. So, without loss of generality, we can assume that $x\neq d$ and c covers b in P (otherwise we start with $\{d, y, x, c\}$ as an N in P). From Theorem 2, d is contained in a minimal cutset K of P. Suppose that K is an antichain and let

$$X=\{U(x)\cap D\}\cup\{D(y)\cap C\} \quad\text{and}\quad Y=\{U(c)\cap C\}\cup\{D(b)\cap D\}.$$

Since C and D are maximal chains and $c\succ b$, $x\succ y$ in P, the chains X and Y are maximal in P. Therefore $K\cap X\neq\emptyset$ and $K\cap Y\neq\emptyset$. Let $u\in K\cap X$ and $v\in K\cap Y$. Since K is an antichain, $v\geqq c$ and $u\leqq a$. Thus $v<u$. This contradiction completes the proof.

Proof of Theorem 4. Suppose that P is an ordered set which contains a generalized $N=(C_1, C_2, A_1, A_2)$. Let A be a maximal antichain in P, containing $A_1\cup A_2$ and let C be a maximal chain in P such that $C_1\cup C_2\subseteq C$. Since $C_1<C_2$ and there is no x such that $C_1<\{x\}<C_2$, for every element y in $C-(C_1\cup C_2)$, either $y>c_2$ for some element c_2 in C_2 or $y<c_1$ for some element c_1 in C_1. But A_2 is coinitial in C_2 and A_1 is cofinal in C_1, thus either $y>a_2$ for some a_2 in A_2 or $y<a_1$ for some a_1 in A_1. Therefore $y\notin A$ and $C\cap A=\emptyset$, which contradicts that A is an antichain cutset in P.

To prove the converse assume that P contains a maximal antichain A and a maximal chain C such that $A\cap C=\emptyset$. Let

$$C_1=\{x\in C\,|\,x<a \text{ for some } a \text{ in } A\}$$

and let

$$C_2=\{x\in C\,|\,x>a \text{ for some } a \text{ in } A\}.$$

Since $A\cap C=\emptyset$ and A is a maximal antichain, $C_1\cap C_2=\emptyset$ and $C_1\cup C_2=C$. If $C_1=\emptyset$ then $(\emptyset, C_2, \emptyset, A)$ is a generalized N, and the dual argument applies if $C_2=\emptyset$. So, we assume that $C_1\neq\emptyset\neq C_2$.

Let α and β be ordinals such that $\alpha=\mathrm{cf}(C_1)$ and $\beta=\mathrm{ci}(C_2)$. (The *cofinality* of a chain C of order type γ, denoted by $\mathrm{cf}(C)$ or $\mathrm{cf}(\gamma)$ too, is the least ordinal α such that there is a subchain C' of C of order type α and cofinal in C. The *coinitiality* of a chain C of order type γ, denoted by $\mathrm{ci}(C)$ or $\mathrm{ci}(\gamma)$ too, is the least ordinal β such that there is a subchain C' of C of order type β^d, the dual of β, and coinitial in C.) Let

$$F=\{x_0<x_1<\ldots<x_i<\ldots\}_{i<\alpha}$$

be a cofinal subset of C_1 and

$$I=\{y_0>y_1>\ldots>y_i\ldots\}_{i<\beta}$$

be a coinitial subset of C_2. Now, we construct, simultaneously, the antichains $(A_1^i)_{i<\gamma}$ and $(A_2^i)_{i<\gamma}$ in P, where $\gamma=\min(\alpha, \beta)$, as follows.

Let a_0 be in A such that $a_0>x_0$, and let b_0 in $A-\{a_0\}$ such that $b_0<y_0$. We set $A_1^0=\{a_0\}$ and $A_2^0=\{b_0\}$.

Suppose we have already constructed $(A_1^i)_{i<\delta}$ and $(A_2^i)_\delta$. If $\delta=\delta'+1$, an isolated ordinal, then let $a_\delta \in A-(A_1^{\delta'}\cup A_2^{\delta'})$ such that $a_\delta>x_\delta$. And let $b_\delta\neq a_\delta$ in $A-(A_1^{\delta'}\cup A_2^{\delta'})$ such that $b_\delta<y_\delta$. We set

$$A_1^\delta = A_1^{\delta'}\cup\{a_\delta\} \quad \text{and} \quad A_2^\delta = A_2^{\delta'}\cup\{b_\delta\}.$$

If δ is a limit ordinal then we set

$$A_1^\delta = \bigcap_{i<\delta} A_1^i \quad \text{and} \quad A_2^\delta = \bigcap_{i<\delta} A_2^i.$$

Since the antichain A is cofinal in C_1 and coinitial in C_2, this construction is possible until A_1^γ and A_2^γ. Without loss of generality we can assume that $\gamma=\alpha\leqq\beta$. Let $A_1=A_1^\alpha$ and $A_2=A-A_1$. The four-tuple (C_1, C_2, A_1, A_2) is a generalized N in P.

References

[1] R. P. Dilworth, A decomposition theorem for partially ordered sets, *Ann. of Math.*, **51** (1950), 161—166.

[2] P. A. Grillet, Maximal chains and antichains, *Fund. Math.*, **15** (1969), 157—167.

[3] D. Higgs, A companion to Grillet's theorem on maximal chains and antichains, *Order*, **1** (1985), 371—375.

[4] I. Rival and N. Zaguia, Antichain cutsets, *Order*, **1** (1985), 235—247.

UNIVERSITY OF OTTAWA
DEPARTMENT OF COMPUTER SCIENCE
OTTAWA, K1N 9B4, CANADA

Acta Sci. Math., 51 (1987), 209—215

Homomorphism of distributive lattices as restriction of congruences

E. TAMÁS SCHMIDT

Dedicated to the memory of András Huhn

1. Introduction. Let I be an ideal of a lattice L. Then the map $\varrho: \Theta \to \Theta_I$, restricting a congruence relation Θ to I is a 0 and 1 preserving lattice-homomorphism of the congruence lattice Con L into Con I. G. GRÄTZER and H. LAKSER [1] have proved the converse for finite lattices:

Theorem A. *Let D and E be finite distributive lattices, and let $\varphi: D \to E$ be a* 0 *and* 1 *preserving homomorphism of D into E. Then there exist a finite lattice L, and an ideal I of L, such that there are isomorphisms $\alpha: D \to \mathrm{Con}\, L$, $\beta: E \to \mathrm{Con}\, I$, satisfying $\beta\varphi = \varrho\alpha$, where $\varrho: \Theta \to \Theta_I$ is the restriction of $\Theta \in \mathrm{Con}\, L$ to I.* (See Figure 1.)

$$\begin{tikzcd} D \arrow[r, "\varphi"] \arrow[d, "\alpha"'] & I \arrow[d, "\beta"] \\ \mathrm{Con}\, L \arrow[r, "\varrho"] & \mathrm{Con}\, I \end{tikzcd}$$

Figure 1

The purpose of this paper is twofold. Firstly, we generalize Theorem A for distributive algebraic lattices satisfying the following condition

($*$) for all compact x, $x \vee \bigwedge(x_i \mid i \in I) = \bigwedge(x \vee x_i \mid i \in I)$,

which is a weaker form of the infinite meet distributivity. Secondly, we win a short proof of Theorem A, which uses a construction given in SCHMIDT [3] and [4].

2. Dual Heyting algebras. Let L be a lattice. The dual pseudocomplement of a relative to b is an element $a * b$ of L satisfying $a \vee x \geqq b$ iff $x \geqq a * b$. A *dual Heyting algebra* is a distributive lattice with 1 in which $a * b$ exists for all $a, b \in L$. The subset

Received November 10, 1985, and in revised form December 5, 1986.

of all compact elements of an algebraic lattice A is denoted by $K(A)$. This $K(A)$ is a join-subsemilattice with smallest element 0. A lattice A is called *arithmetic* iff it is algebraic and $K(A)$ is a sublattice of L.

Lemma 1. *Let L be a distributive arithmetic lattice, whose unit element is compact. L satisfies the condition $(*)$ if and only if $K(L)$ is a dual Heyting algebra.*

Proof. First, let $K(L)$ be a dual Heyting algebra. Then L is isomorphic to the lattice of all ideals of $K(L)$, and the compact elements of the ideal lattice are precisely the principal ideals. Therefore we have to show that

$$(x]\vee\bigwedge(J_i\,|\,i\in I)=\bigwedge((x]\vee J_i\,|\,i\in I)$$

where the J_i-s are ideals of $K(L)$. It is enough to verify that the right side is contained in the left side. Let $a\in\bigwedge((x]\vee J_i|\, i\in I)$, then $a\in(x]\vee J_i$ for all $i\in I$, i.e., $a\leqq x\vee y_i$ for suitable $y_i\in J_i$. $K(L)$ is a dual Heyting algebra, therefore $x*a$ exists and $x*a\leqq y_i$ implies $x*a\in J_i$, $i\in I$. Thus $x*a\in\bigwedge(J_i\, i\in I)$. By the definition of $x*a$ we have $a\leqq x\vee(x*a)$, i.e. $a\in(x]\vee\bigwedge(J_i|\, i\in I)$.

By assumption L is a distributive arithmetic lattice with compact unit element, thus $K(L)$ is a bounded distributive lattice. Consider all u_i-s such that $a\vee u_i\geqq b$. Then $b\in\bigwedge((a]\vee(u_i])$. Applying $(*)$ we obtain $b\in(a]\vee\bigwedge(u_i]$, i.e. there exists a $z\in\bigwedge(u_i]$ such that $a\vee z\geqq b$. Obviously $z=a*b$.

By Lemma 1, we can work with dual Heyting algebras, namely L is determined by $K(L)$.

Let L be a $\{0,1\}$-sublattice of the Boolean lattice B. Then L is said to R-generate B if L generates B as a ring. The following lemma is due to H. M. MacNeille (see G. GRÄTZER [2]).

Lemma 2. *Let B be R-generated by L. Then every $a\in B$ can be expressed in the form*

$$a_0+a_1+\ldots+a_{n-1},\quad a_0\leqq a_1\leqq\ldots\leqq a_{n-1},\quad a_0,\ldots,a_{n-1}\in L.$$

A sublattice L' of a dual Heyting algebra L is called a *subalgebra* if for every $x\in L$ there exists a smallest $\bar{x}\in L'$ that $x\leqq\bar{x}$.

Lemma 3. *A subalgebra of a dual Heyting algebra is a dual Heyting algebra.*

Proof. Let L' be a subalgebra of L and let $a,b\in L'$, $a\leqq b$. Then $a*b$ exists in L, and it is easy to verify that $\overline{a*b}$ is the dual pseudocomplement of a relative to b in L'. It is clear that if the dual pseudocomplement exists for comparable pairs then there exists for arbitrary pairs.

For a bounded distributive lattice L we shall denote by $B(L)$ the Boolean lattice R-generated by L.

Lemma 4. *Let L be a dual Heyting algebra. Then L is a subalgebra of $B(L)$.*

Proof. Let L be a dual Heyting algebra. Then by Lemma 2 every x can be expressed in the form $x=a_0+\dots+a_{n-1}$. We prove the existence of $\bar{x}$ by induction on n. If $n=1$, i.e. $x=a_0$ then $x\in L$ hence $\bar{x}=x$. For $n=2$, i.e. $x=a_0+a_1$, x is the relative complement of a_0 in the interval $[0, a_1]$. Then $a_0\vee y\geqq a_1$ and $y\in L$ imply $y\geqq a_0+a_1$, thus $\overline{a_0+a_1}$ exists and $\overline{a_0+b_1}=a_0*a_1$. Let us assume that $n>2$. $a_{n-2}+a_{n-1}$ is the relative complement of a_{n-2} in the interval $[0, a_{n-1}]$, hence $a_{n-3}\wedge(a_{n-2}+a_{n-1})\leqq a_{n-2}\wedge(a_{n-2}+a_{n-1})=0$. Obviously $a_0+\dots+a_{n-3}\leqq a_{n-2}$, thus $(a_0+\dots+a_{n-3})\wedge(a_{n-2}+a_{n-1})=0$. This implies that $a_0+\dots+a_{n-1}=(a_0+\dots+a_{n-3})+$ $+(a_{n-2}+a_{n-1})=(a_0+\dots+a_{n-3})\vee(a_{n-2}+a_{n-1})$.

Let x, y be arbitrary elements of $B(L)$ such that $\bar{x}$ and $\bar{y}$ exist. We prove that $\overline{x\vee y}$ exists and $\overline{x\vee y}=\bar{x}\vee\bar{y}$. Let $x\vee y\leqq z\leqq\bar{x}\vee\bar{y}$, $z\in L$. Then we get from $x, y\leqq x\vee y$ that $x\leqq\bar{x}\wedge z\in L$, $y\leqq\bar{y}\wedge z\in L$, and we conclude that $\bar{x}\leqq z$, $\bar{y}\leqq z$, i.e. $z=\bar{x}\vee\bar{y}$, which proves $\overline{x\vee y}=\bar{x}\vee\bar{y}$. Applying this equality for $x=a_0+\dots+a_{n-3}$ and $y=$ $=a_{n-2}+a_{n-1}$ we obtain that $\overline{x\vee y}=\overline{x+y}$ exists.

Lemma 5. *Let Θ be a compact congruence relation of a dual Heyting algebra L. Then L/Θ is a dual Heyting algebra.*

Proof. The compact congruence relations are exactly the finite joins of principal congruence relations. To prove the lemma, by the Second Isomorphism Theorem we may assume that Θ is a principal congruence relation, i.e. $\Theta=\Theta(u, v)$, $u\leqq v$.

Let L be a dual Heyting algebra. We prove that each congruence class of $\Theta(u, v)$ contains a smallest element. In distributive lattices $\Theta(u, v)$ has the following description (see [2], p. 74): $a\equiv b\ \big(\Theta(u, v)\big)$ iff $v\vee a=v\vee b$ and $u\wedge a=u\wedge b$. Let b be a fixed element of L. Then $v\vee a=v\vee b$ implies that $a\geqq v*(v\vee b)$. Therefore $v*(v\vee b)$ is the least element of the $\Theta(u, v)$-class containing b. Now, let $a<b$ and let c denote the least element of the $\Theta(u, v)$-class containing b. Let $[x]$ denote the $\Theta(u, v)$-class containing x. Then obviously $[a]*[b]=[a*c]$.

Corollary. *Every Θ-class of a compact congruence relation Θ of a dual Heyting algebra contains a smallest element.*

3. The main theorem. In this section we formulate our main theorem and then we give two special representations of dual Heyting algebras.

Theorem B. *Let D and E be dual Heyting algebras, and let $\varphi\colon D\to E$ be a 0 and 1 preserving homomorphism of D into E such that the congruence kernel* Ker φ *is a compact congruence relation and $F=\operatorname{Im}\varphi$ is a subalgebra of E. Then there exist a lattice L, and a principal ideal I of L, such that there are isomorphisms $\alpha\colon D\to K(\operatorname{Con} L)$, $\beta\colon E\to J(\operatorname{Con} I)$ satisfying $\beta\varphi=\varrho\alpha$, where $\varrho\colon\Theta\to\Theta_I$ is the restriction of $\Theta\in\operatorname{Con} L$ to I.*

If L_1 and L_2 are lattices with zero elements 0_1 resp. 0_2 then in the direct product $L_1 \times L_2$ the elements $\langle x, 0_2 \rangle$ $(x \in L_1)$ form an ideal L_1' isomorphic to L_1. Therefore we can identify L_1 with L_1' and similarly L_2 with the ideal $L_2' = \{\langle 0_1, x \rangle\}$.

Let Θ be the congruence kernel of the homomorphism $\varphi \colon D \to E$. By our assumption Θ is a compact congruence relation of D. On the other hand D is a bounded distributive lattice, therefore the unit of Con D is compact. The compact elements of Con D form a Boolean lattice (see [2], p. 86, Exercise 41), consequently Θ has a complement Θ' in Con D. Then D is a subdirect product of D/Θ and D/Θ', therefore $D \subseteq D/\Theta \times D/\Theta'$.

$F = \mathrm{Im}\, \varphi$ is a $\{0, 1\}$-sublattice of E and F is isomorphic to D/Θ; we identify D/Θ and F. Hence we may consider D as a $\{0, 1\}$-sublattice of $E \times D/\Theta'$. Let e be the unit of E and $\pi_1(x) = x \wedge e$ denotes the projection map of $E \times D/\Theta'$ onto E. Observe, that the restriction of π_1 to D $(\subseteq E \times D/\Theta')$ gives the homomorphism φ (see Figure 2).

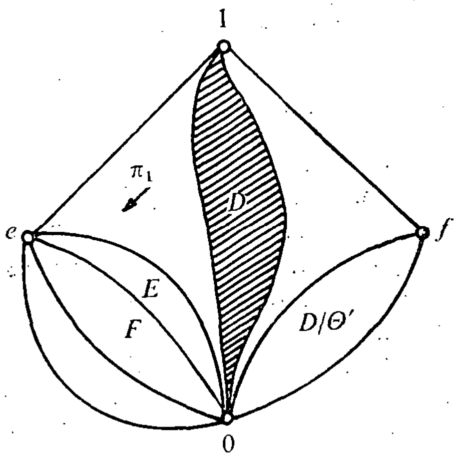


Figure 2

Lemma 6. *$E \times D/\Theta'$ is a dual Heyting algebra and D is a subalgebra of $E \times D/\Theta'$.*

Proof. By our assumptions D and E are dual Heyting algebras and Θ' is a compact congruence relation. Hence by Lemma 5 D/Θ' and thus $E \times D/\Theta'$ are dual Heyting algebras. $F = \mathrm{Im}\, \varphi$ is a subalgebra of E, hence by Lemma 3 F and $F \times D/\Theta'$ are dual Heyting algebras.

We have seen that D is a subdirect product of F and D/Θ'. First we show that D is a subalgebra of the dual Heyting algebra $F \times D/\Theta'$. An arbitrary element of $F \times D/\Theta'$ can be written in the form $x = \langle [a]\Theta, [b]\Theta' \rangle$ where $a, b \in D$. By Corollary of Lemma 5 the congruence classes of Θ and Θ' have smallest elements. Let now a_0 and b_0 be the smallest elements of $[a]\Theta$ resp. $[b]\Theta'$. Then $\langle[a_0\wedge b_0]\Theta, [a_0\wedge b_0]\Theta'\rangle\in D$. Obviously this element is in D, the smallest one which is greater or equal than x, i.e. $\bar{x}$ exist. This proves that D is a subalgebra of $F\times D/\Theta'$. On the other hand F is a subalgebra of E, consequently $F\times D/\Theta'$ is a subalgebra of $E\times D/\Theta'$, which proves finally that D is a subalgebra of $E\times D/\Theta'$ (namely a subalgebra of a subalgebra is again a subalgebra).

In [3] (or see [4]) there was given a special lattice construction to prove that the lattice of all ideals of a dual Heyting algebra is isomorphic to the congruence lattice of some lattice. The most important properties of this construction are summarized in the following lemma.

Lemma 7. *Let K be a $\{0, 1\}$-subalgebra of a Boolean lattice A, and let $\varepsilon\colon K\to A$ be the identy map. There exists a bounded lattice M with the following properties:*

(i) *M contains three elements u, v, w such that $\{0, u, v, w, 1\}$ form a sublattice isomorphic to the diamond M_3. There are isomorphisms $\mu\colon (u]\to K$ and $\tau\colon (v]\to A$. If for $x\in K$ $\pi(x)$ means $(x\vee w)\wedge v$ then $\tau\pi=\varepsilon\mu$.*

(ii) *The map $x\to x\vee u$ $(x\leqq v)$ is an isomorphism of $(v]$ onto the filter $[u)$.*

(iii) *A congruence relation $\Theta(0, x)$ of $(v]$ can be extended to M iff $\tau(x)\in\varepsilon(K)$, and every compact congruence relation of M is the extension of a congruence relation $\Theta(0, x)\in\mathrm{Con}\,(v]$.*

Remark. A is a Boolean lattice, therefore every compact congruence relation of A $(\cong(v])$ can be written in the form $\Theta(0, x)$. Condition (iii) implies that $\mathrm{Con}\, M\cong$ $\cong I(K)$, i.e. $K(\mathrm{Con}\, M)\cong K$.

4. The proof of Theorem B. We apply Lemma 7 twice to get two lattices M_1 and M_2. Then we use the so called Hall—Dilworth gluing construction which yields a lattice L having the properties required in the theorem.

By Lemma 6 D is a subalgebra of the dual Heyting algebra $E\times D/\Theta'$ and by Lemma 4 $E\times D/\Theta'$ is a subalgebra of $B(E\times D/\Theta')$. Consequently D is a subalgebra of $B(E\times D/\Theta')$. Then we can choose in Lemma 7 $K=D$ and $A=B(E\times D/\Theta')$. We obtain the lattice M_1 with a diamond $\{0_1, u_1, v_1, w_1, 1_1\}$ given in condition (i) of Lemma 7. In the second case we consider $K=E\times B(D/\Theta')$ and $A=B(E\times D/\Theta')$. By Lemma 4 E is a subalgebra of $B(E)$ hence $E\times B(D/\Theta')$ is a subalgebra of $B(E)\times$ $\times B(D/\Theta')=B(E\times D/\Theta')$. The resulting lattice is M_2 with the diamond $\{0_2, u_2, v_2, w_2, 1_2\}$.

By condition (i) of Lemma 7 the ideal $(v_1]$ of M_1 is isomorphic to $B(E\times D/\Theta')$. On the other hand by condition (ii) the filter $[u_2)$ of M_2 is isomorphic to $B(E\times D/\Theta')$. Consequently we have an isomorphism $\delta\colon [u_2)\to(v_1]$. We apply the Hall—Dilworth gluing construction which gives a lattice L having M_1 as a filter and M_2 as an ideal.

(L is the set of all $x \in M_1 \cup M_2$, we identify x with $\delta(x)$ for all $x \in [u_2)$; $x \leqq y$ has unchanged meaning if $x, y \in M_1$ or $x, y \in M_2$ and $x < y$, $x, y \notin [u_2) = (v_1]$ iff $x \in M_2$, $y \in M_1$ and there exists a $z \in [u_2)$ such that $x < z$ in M_2 and $z < y$ in M_1.) The lattice L is given by Figure 3 where $B = B(E \times D/\Theta')$:

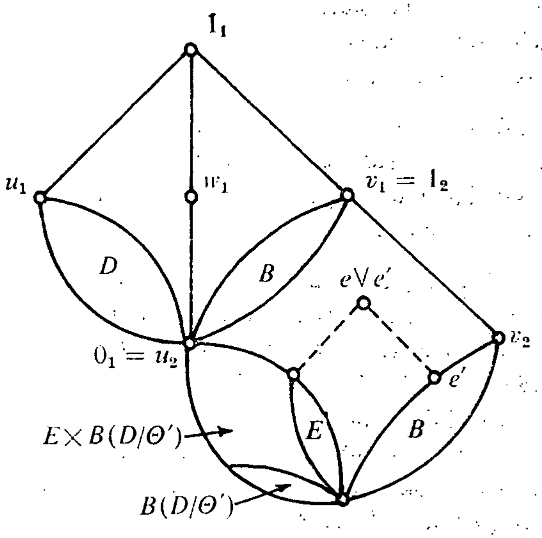


Figure 3

The function $\pi(x) = (x \vee w_2) \wedge v_2$ yields the element $e' = \pi(e) = (e \vee w_2) \wedge v_2 \leqq v_2$. Let I be the principal ideal generated by $e \vee e'$. We have to prove that the pair L, I satisfies the properties given in Theorem B.

(1) First we prove that $\operatorname{Con} L \cong I(D)$ i.e. D is isomorphic to the semilattice of all compact congruences of L. Every congruence relation Θ of L is determined by its restrictions Θ_{M_1} and Θ_{M_2} to M_1 resp. M_2. By condition (iii) of Lemma 7 Θ_{M_1} is determined by its restriction to $(v_1]$ and similarly Θ_{M_2} is determined by its restriction to $(v_2]$. But the interval $[0_1, v_1]$ is a transpose of $[0_2, v_2]$, hence we get that Θ is determined by its restriction to the ideal $(v_2]$. This ideal is a Boolean lattice, thus every compact congruence relation of $(v_2]$ has the form $\Theta(0_2, x)$, $x \in (v_2]$. Let now, $\Theta(0_2, x)$ be a congruence relation of $(v_2]$. Under what conditions for x has this congruence relation an extension to L? Condition (iii) of Lemma 7 gives the following isomorphisms:

$$\text{in } M_1,\quad \mu_1\colon (u_1] \to D,\quad \tau_1\colon (v_1] \to B(E \times D/\Theta'),$$
$$\text{in } M_2,\quad \mu_2\colon (u_2] \to E \times B(D/\Theta'),\quad \tau_2\colon (v_2] \to B(E \times D/\Theta').$$

If $\varepsilon_1\colon D \to B(E \times D/\Theta')$ and $\varepsilon_2\colon E \times B(D/\Theta')$ denote the identity maps, then $\tau_1\pi_1 = \varepsilon_1\mu_1$ and $\tau_2\pi_2 = \varepsilon_2\mu_2$ where $\pi_i(x) = (x \vee w_i) \wedge v_i$ $(i = 1, 2)$. By condition (iii) of

Lemma 7, the congruence relation $\Theta(0_2, x)$ can be extended to M_2 iff $\tau_2(x)\in \in\varepsilon_2(E\times B(D/\Theta'))$. Similarly, in M_1 we get that the congruence relation $\Theta(0_1, 0_1\vee x)$ of $(v_1]$ can be extended to M_1 iff $\tau_1(0_1\vee x)\in\varepsilon_1(D)$. Obviously the minimal extensions of $\Theta(0_2, x)$ and $\Theta(0_1, 0_1\vee x)$ to L are the same and $\varepsilon_1(D)$ is a sublattice of $\varepsilon_2(E\times B(D/\Theta'))$, so we obtain that $\Theta(0_2, x)$ has an extension to L iff $\tau_1(0_1\vee x)\in\varepsilon_1(D)$. This proves $\mathrm{Con}\ L\cong I(D)$.

(2) Secondly we show $\mathrm{Con}\ I\cong I(E)$. E is a direct factor of $(u_2]$ and $(e']$ is isomorphic to $B(E)$. Obviously $B(E\times B(D/\Theta'))=B(E)\times B(D/\Theta')$ hence the principal ideal $I=(e\vee e']$ is a direct factor of M_2. This means that I is again a lattice given by Lemma 7, namely if $K=E$ and $A=B(E)$. Thus by condition (iii) we have $K(\mathrm{Con}\ I)\cong E$, i.e. $\mathrm{Con}\ I\cong I(E)$.

(3) Finally, let Θ be a compact congruence relation of L. We have seen that Θ is the extension of some $\Theta(0_2, x)\in\mathrm{Con}\ (u_2]$ where $\tau_2(x)\in\varepsilon_1(D)$, i.e. the restriction $\Theta\to\Theta_I$ is determined by the projection $D\to E\times D/\Theta'$. As we have seen this is exactly the given homomorphism φ.

References

[1] G. Grätzer and H. Lakser, Homomorphisms of distributive lattices as restrictions of congruences, *Canad. J. Math.*, **38** (1986), 1122—1134.

[2] G. Grätzer, *General lattice theory*, Akademie-Verlag (Berlin, 1978).

[3] E. T. Schmidt, Zur Charakterisierung der Kongruenzverbände der Verbände, *Math. Časopis*, **18** (1968), 3—20.

[4] E. T. Schmidt, *A Survey on Congruence Lattice Representations*, Teubner-Texte zur Mathematik, Band 42, Teubner (Leipzig, 1982).

MATHEMATICAL INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES
P. O. BOX 127
1364 BUDAPEST, HUNGARY

Acta Sci. Math., 51 (1987), 217—219

A characterization of semimodularity in lattices of finite length

M. STERN

Dedicated to the memory of András P. Huhn

1. Introduction. In this note we consider only lattices of finite length. By $x \prec y$ we mean that x is a lower cover of y. If L is a lattice of finite length, we denote by $J(L)$ the set of all join-irreducible elements ($\neq 0$) of L. Equivalently, an element u is in $J(L)$ if and only if it has precisely one lower cover which will be denoted by u'. A lattice L of finite length is called (upper) semimodular if the neighbourhood condition (N) holds in L:

$$\text{(N)} \qquad a \wedge b \prec a \Rightarrow b \prec a \vee b \quad (a, b \in L).$$

It is the aim of this paper to show that (N) can be replaced by a seemingly weaker condition $(\overline{\mathrm{N}})$ which may be called restricted neighbourhood condition or neighbourhood condition for join-irreducible elements:

$$(\overline{\mathrm{N}}) \qquad u \wedge b = u' \prec u \Rightarrow b \prec u \vee b \quad (u \in J(L),\ b \in L).$$

After a preliminary lemma in Section 2, we show the equivalence of (N) and $(\overline{\mathrm{N}})$ in Section 3. Applying this result to the atomistic case (i.e., to the case in which each join-irreducible element ($\neq 0$) is an atom) we get the well-known result that in these lattices semimodularity is equivalent to the so-called covering property.

2. Preliminary remarks. In this section we prove the following

Lemma. *Let L be a lattice of finite length. If $c \prec d$ $(c, d \in L)$, then there exists a join-irreducible element $u \in J(L)$ such that $u \leqq d$, $u \not\leqq c$ and $u \wedge c = u'$.*

Proof. If $d \in J(L)$, then put $u = d$. Let now $d \notin J(L)$ and consider the set of all $v \in J(L)$ which have the property $v < d$ and $v \not\leqq c$. It is clear that this set is not empty. Choose an element $u \in J(L)$ which is minimal with respect to this property. Since L is of finite length, such a minimal element always exists. From $u < d$

Received March 1, 1985, and in revised form November 26, 1985.

and $u \not\leqq c$ it follows that $u \wedge c \leqq u'$. We show that equality holds. From the assumption $u \wedge c < u'$ we get the existence of an element $u_* \in J(L)$ having the properties $u_* \leqq u'$ and $u_* \not\leqq u \wedge c$. This implies

$$u_* \leqq u' \prec u < d \quad \text{and} \quad u_* \not\leqq c.$$

(Note that $u_* \leqq c$ yields together with $u_* < u$ that $u_* \leqq u \wedge c$, a contradiction.) But this in turn contradicts the minimality of $u \in J(L)$. Thus our assumption was false, i.e., we have $u \wedge c = u'$, which was to be proved.

Remark. The preceding lemma was implicitly used in the proof of the main theorem of [2] and it was explicitly given in [3]. We have included the proof here in order to make the paper self-contained. This lemma is a generalization of a property which is trivially fulfilled in atomistic lattices of finite length.

3. Results. Using the lemma of the preceding section, we prove here the following

Theorem. *Let L be a lattice of finite length. Then the neighbourhood condition* (N) *holds if and only if the restricted neighbourhood condition* ($\overline{\text{N}}$) *holds.*

Proof. (N)$\Rightarrow$($\overline{\text{N}}$): This implication is obviously true.

($\overline{\text{N}}$)$\Rightarrow$(N): Assuming ($\overline{\text{N}}$) we show that (N) also holds. In other words, in lattices of finite length the restricted neighbourhood condition already implies the (upper) semimodularity of the lattice. Without loss of generality we may assume that $a, b \in L$ are incomparable elements. In order to prove the assertion assume

$$(*) \qquad a \wedge b \prec a \quad (a, b \in L).$$

We show that then $b \prec a \vee b$ also holds. If $a = u \in J(L)$, it follows by ($\overline{\text{N}}$) that $b \prec u \vee b = a \vee b$ and nothing is to be proved. Assume now $a \notin J(L)$. By the lemma of Section 2 there exists a join-irreducible element $u \in J(L)$ having the properties $u < a$, $u \not\leqq a \wedge b$ and $u' = u \wedge (a \wedge b) = u \wedge b$. From $u < a$, $u \not\leqq a \wedge b$ and $a \wedge b \prec a$ we obtain $a = (a \wedge b) \vee u$. This means that

$$a \vee b = (a \wedge b) \vee u \vee b = b \vee u.$$

Now

$$u' = u \wedge b \prec u$$

implies by ($\overline{\text{N}}$) that

$$(**) \qquad b \prec u \vee b = a \vee b.$$

To sum it up: under the assumption ($\overline{\text{N}}$), the relation $(*)$ implies $(**)$ which means that the lattice is semimodular. This finishes the proof of the theorem.

Corollary. *Let L be an atomistic lattice of finite length. Then L is (upper) semimodular if and only if the covering property*

(C) $$p(\in L) \text{ atom},\ b\in L,\ b\wedge p=0 \Rightarrow b\prec b\vee p$$

holds.

Proof. In the atomistic case ($\bar{\mathrm{N}}$) reduces to (C) implying semimodularity by the preceding theorem. The converse statement is obviously true.

We remark that the assertion of the corollary holds even for arbitrary atomistic lattices by [1, Theorem 7.10, p. 32].

Acknowledgments. The author thanks the referee for his remarks according to which the paper was condensed. The original version was rewritten during a stay at Asmara University (Ethiopia) in the Academic Year 1985/86. The author thanks the Ethiopian colleagues for their support.

References

[1] F. Maeda—Sh. Maeda, *Theory of Symmetric Lattices*, Springer-Verlag (Berlin—Heidelberg—New York, 1970).

[2] M. Stern, Semimodularity in lattices of finite length, *Discrete Mathematics*, **41** (1982), 287—293.

[3] M. Stern, Exchange properties in lattices of finite length, *Wiss. Z. Univ. Halle*, **31** (1982), M. H. 3, 15—26.

MARTIN-LUTHER-UNIVERSITÄT
SEKTION MATHEMATIK
UNIVERSITÄTSPLATZ 6
402 HALLE, GDR

Acta Sci. Math., 51 (1987), 221—227

Triply transitive algebras

LÁSZLÓ SZABÓ

To the memory of András Huhn

In [7] P. SCHOFIELD proved that if G is a triply transitive permutation group on an at least four element finite set M and f is a surjective operation on M depending on at least two variables then the clone F generated by $G\cup\{f\}$ either equals the set of all operations on M or $F\subseteq L$ where L is a maximal clone of quasilinear operations on M. The aim of this paper is to improve this result by proving that the inclusion $F\subseteq L$ is actually an equality (Theorem 8).

In [6] R. PÖSCHEL described all finite relationally incomplete homogeneous relation algebras. As an application of our theorem we also improve this result by giving all at least four element finite relationally incomplete relation algebras having triply transitive automorphism groups (Theorem 9).

2. Preliminaries

Let M be a nonempty set. The set of all n-ary operations on M will be denoted by $O_M^{(n)}$ $(n\geqq 1)$, and we set $O_M=\bigcup_{n\geqq 1} O_M^{(n)}$. An operation $f\in O_M$ is *idempotent* if for every $a\in M$ we have $f(a, ..., a)=a$; f is *nontrivial* if it is not a projection. If f depends on at least two variables and takes on all values from M then it is called *essential.*

For $h\geqq 1$ the set of h-ary relations on M (i.e. subsets of M^h) will be denoted by $R_M^{(h)}$; furthermore we set $R_M=\bigcup_{h\geqq 1} R_M^{(h)}$. An operation $f\in O_M^{(n)}$ is said to *preserve* a relation $\varrho\in R_M^{(h)}$ if ϱ is a subalgebra of the h-th direct power of the algebra $\langle M; f\rangle$. For $R\subseteq R_M$ the symbol Pol R denotes the set of all operations from O_M preserving

Received June 24, 1986, and in revised form October 27, 1986.

Research partially supported by Hungarian National Foundation for Scientific Research grant no. 1813.

each relation in R, and for $F\subseteq O_M$ the symbol Inv F denotes the set of all relations from R_M preserved by each operation in F. The correspondences $R\to\operatorname{Pol} R$ and $F\to\operatorname{Inv} F$ establish a Galois connection between the subsets of R_M and the subsets of O_M. For $F\subseteq O_M$ and $R\subseteq R_M$ we set $\langle F\rangle=\operatorname{Pol}\operatorname{Inv} F$ and $[R]=\operatorname{Inv}\operatorname{Pol} R$.

By a *clone of operations* on M we mean a subset $F\subseteq O_M$ which contains the projections and is closed with respect to superposition. It is known (cf. e.g. [5]) that, for finite M, a subset $F\subseteq O_M$ is a clone if and only if $F=\langle F\rangle$. By a *clone of relations* we mean a subset $R\subseteq R_M$ satisfying the equality $R=[R]$. We remark that for finite M there exists also an internal definition for $[R]$, namely $[R]$ is the set of all relations which are definable by a first order formula in which only $\exists$, $\wedge$, $=$, and relations (i.e. predicates) of R occur. For more details cf. [5].

By a *relation algebra* on the set M we mean a pair $\langle M; R\rangle$ where $R\subseteq R_M$. We say that $\langle M; R\rangle$ is nontrivial if $\operatorname{Pol} R\neq O_M$. A permutation π on M is an *automorphism* of $\langle M; R\rangle$ if $\varrho\pi\subseteq\varrho$ and $\varrho\pi^{-1}\subseteq\varrho$ for every $\varrho\in R$. The symbol Aut $\langle M; R\rangle$ denotes the group of all automorphisms of $\langle M; R\rangle$.

If f is an n-ary operation on M then $f^{\bullet}$ denotes the $(n+1)$-ary relation $\{(a_1, \dots, a_n, f(a_1, \dots, a_n))\mid a_1, \dots, a_n\in M\}$. Two relation algebras $\langle M; R_1\rangle$ and $\langle M; R_2\rangle$ are *equivalent* if $[R_1]=[R_2]$.

If $n\geqq 1$ and q is a prime power then $V(n, q)$ denotes the n-dimensional vector space over the field $GF(q)$. In this note by a *linear operation* over $V(n, q)$ we mean an operation of the form $\sum_{i=1}^{m} x_iA_i+v$ where $v\in V(n, q)$ and the A_i $(1\leqq i\leqq m)$ are linear transformations of $V(n, q)$. Clearly, such an operation depends on its i-th variable if and only if $A_i\neq 0$, and is surjective if and only if $V(n, q)$ is spanned by its subspaces $\operatorname{Im} A_i$, $i=1, \dots, m$. The set of all linear operations over $V(n, q)$ will be denoted by $ACL(n, q)$; and as usual $AGL(n, q)$ resp. $GL(n, q)$ denote the set of all linear permutations resp. the set of all linear permutations fixing the zero vector $0\in V(n, q)$.

Let us denote by $\mathscr{A}_n$ $(n\geqq 1)$ the alternating group of degree n. It is well known (see e.g. [3]) that $GL(4, 2)\cong\mathscr{A}_8$, and thus $GL(4, 2)$ contains subgroups isomorphic to $\mathscr{A}_7$.

We need the following results.

Proposition 1 ([3], [4]). *If G is a subgroup of $GL(4, 2)$ and $G\cong\mathscr{A}_7$ then G is doubly transitive on $V(4, 2)\setminus\{0\}$, moreover, for any two triples u_1, u_2, u_3 and v_1, v_2, v_3 of linearly independent vectors in $V(4, 2)$ there is exactly one permutation $A\in G$ such that $u_iA=v_i$, $i=1, 2, 3$. Consequently, if T is the group of all translations on $V(4, 2)$ then $G\ltimes T$ is a triply transitive proper subgroup of $AGL(4, 2)$.*

Consider the elements of $GL(4, 2)$ as 4×4 matrices over $GF(2)$ in a fixed basis of $V(4, 2)$. Let G be a subgroup of $GL(4, 2)$ with $G\cong\mathscr{A}_7$. Consider the subgroup

G^* of $GL(4, 2)$, given by $G^* = \{A^* \mid A \in G\}$ where A^* is the transpose of A. Then clearly $G^* \cong \mathscr{A}_7$. Combining this fact with Proposition 1 we immediately get the following statement.

Proposition 2. *Let G be a subgroup of $GL(4, 2)$ with $G \cong \mathscr{A}_7$, and consider the elements of $GL(4, 2)$ as 4×4 matrices over $GF(2)$ in a fixed basis of $V(4, 2)$. Then for any numbers $1 \leqq i_1 < i_2 < i_3 \leqq 4$ and for any linearly independent 4-dimensional row (column) vectors $u_{i_1}, u_{i_2}, u_{i_3}$ over $GF(2)$ there is exactly one element $A \in G$ such that the i-th row (column) of A coincides with u_i for $i = i_1, i_2, i_3$.*

Theorem A (Cameron and Kantor [1]). *If H is a triply transitive proper subgroup of $AGL(n, 2)$ then $n = 4$ and H is $\mathscr{A}_7 \rtimes T$ in $AGL(4, 2)$. Moreover, if G is a doubly transitive proper subgroup of $GL(n, 2)$ (on $V(n, 2) \setminus \{0\}$) then $n = 4$ and G is $\mathscr{A}_7$ in $GL(4, 2)$.*

Theorem B (Szabó and Szendrei [9]). *If $|V(n, q)| \geqq 3$ then $\langle AGL(n, q) \cup \cup \{f\} \rangle = ACL(n, q)$ for every essential operation $f \in ACL(n, q)$.*

Theorem C (Schofield [7]). *If M is a finite set, $|M| \geqq 4$, G is a triply transitive permutation group on M and $f \in O_M$ is an essential operation, then either $\langle G \cup \{f\} \rangle = O_M$ or $|M| = 2^n$ for some $n \geqq 2$ and $\langle G \cup \{f\} \rangle \subseteqq ACL(n, 2)$.*

3. Lemmas

In this section we give some preparatory lemmas.

Lemma 3 (Schofield [7]). *If H is a triply transitive permutation group and f is an essential operation on an at least four element finite set M then $\langle H \cup \{f\} \rangle$ contains all constant operations and an operation taking on m values for some m with $2 \leqq m < |M|$.*

From now on in this section let G denote a subgroup of $GL(4, 2)$ isomorphic to $\mathscr{A}_7$, and let A, A_1, A_2 be unary linear operations on $V(4, 2)$ fixing the zero vector 0. For any unary linear operation X fixing 0, the symbol $G(X)$ denotes the set of all unary linear operations generated by $G \cup \{X\}$.

Lemma 4. *If $\operatorname{Im} A \neq V(4, 2)$, then there is a B in G such that $\operatorname{Im} BA = \operatorname{Im} A$ and $(BA)^2 = BA$.*

Proof. Let $\dim(\operatorname{Im} A) = n \ (\leqq 3)$ and let $u_1, \ldots, u_n$ be a basis of $\operatorname{Im} A$. Choose elements $v_1, \ldots, v_n \in V(4, 2)$ such that $v_i A = u_i$, $i = 1, \ldots, n$. It is easy to see that $v_1, \ldots, v_n$ are linearly independent, and therefore, by Proposition 1, there is a $B \in G$ such that $u_i B = v_i$ $(i = 1, \ldots, n)$. Then $u_i BA = u_i$ $(i = 1, \ldots, n)$ showing that $\operatorname{Im} BA = \operatorname{Im} A$ and $(BA)^2 = BA$.

Lemma 5. *Suppose* $A^2=A$, $\operatorname{Im} A\neq V(4,2)$, *and let* U *be a proper subspace of* $\operatorname{Im} A$ *with* $|U|\geqq 2$. *Then there is a* $B\in G(A)$ *such that* $\operatorname{Im} BA=U$ *and* $(BA)^2=BA$.

Proof. First consider the case when $\dim(\operatorname{Im} A)=3$ and $\dim U=2$. Let u_1, u_2, u_3, u_4 be a basis of $V(4,2)$ such that u_1, u_2 and u_1, u_2, u_3 are bases of U and $\operatorname{Im} A$, respectively, and $u_4\in\operatorname{Ker} A$. By Proposition 1, there is a $C\in G$ such that $u_1C=u_1$, $u_2C=u_2$ and $u_3C=u_4$. Then we have $u_1ACA=u_1$, $u_2ACA=u_2$, $u_3ACA=0$ and $u_4ACA=0$. Therefore if $AC=B$ then $\operatorname{Im} BA=U$ and $(BA)^2=BA$.

Now suppose that $\dim(\operatorname{Im} A)=2$ and $\dim U=1$. Choose a basis u_1, u_2, u_3, u_4 of $V(4,2)$ such that u_1 and u_1, u_2 are bases of U and $\operatorname{Im} A$, respectively, and $u_3, u_4\in\operatorname{Ker} A$. Again by Proposition 1, there is a $C\in G$ such that $u_1C=u_1$ and $u_2C=u_3$. Now if $B=AC$ then we have $\operatorname{Im} BA=U$ and $(BA)^2=BA$.

Finally the statement in the case $\dim(\operatorname{Im} A)=3$ and $\dim U=1$ follows from the previous two cases.

Lemma 6. *If* $\operatorname{Im} A\neq V(4,2)$, *and* U *is a subspace of* $V(4,2)$ *such that* $\dim U=\dim(\operatorname{Ker} A)$ *then there is a* $B\in G$ *such that* $\operatorname{Im} BA=\operatorname{Im} A$ *and* $\operatorname{Ker} BA=U$.

Proof. Let $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ be bases of U and $\operatorname{Ker} A$, respectively. Since $1\leqq n\leqq 3$, by Proposition 1 there is a $B\in G$ such that $u_iB=v_i$, $i=1, \ldots, n$. Then $\operatorname{Im} BA=\operatorname{Im} A$ and $\operatorname{Ker} BA=U$.

Lemma 7. *Suppose that* $\operatorname{Im} A_1, \operatorname{Im} A_2\neq V(4,2)$, *and* $\operatorname{Im} A_1\not\subseteq\operatorname{Im} A_2$, $\operatorname{Im} A_2\not\subseteq\operatorname{Im} A_1$. *Then there are* $B_1\in G(A_1)$ *and* $B_2\in G(A_2)$ *such that* $\operatorname{Im}(B_1A_1+B_2A_2)=\operatorname{Im} A_1+\operatorname{Im} A_2$.

Proof. Let $U_1\subseteqq\operatorname{Im} A_1$ and $U_2\subseteqq\operatorname{Im} A_2$ be subspaces such that $U_1\cap U_2=\{0\}$ and $U_1+U_2=\operatorname{Im} A_1+\operatorname{Im} A_2$. Then applying Lemmas 4 and 5 we get $C_1\in G(A_1)$ and $C_2\in G(A_2)$ such that $\operatorname{Im} C_iA_i=U_i$ and $(C_iA_i)^2=C_iA_i$, $i=1,2$. Since $U_1\cap U_2=\{0\}$, we have $\dim U_1+\dim U_2\leqq 4$. Therefore $\dim(\operatorname{Ker} C_1A_1)\geqq\dim U_2$. Now, by Lemma 6, there is a $D_1\in G$ such that $\operatorname{Im} D_1C_1A_1=U_1$ and $\operatorname{Ker} D_1C_1A_1\supseteqq U_2$. If we choose $B_1=D_1C_1$ and $B_2=C_2$, then we have $\operatorname{Im}(B_1A_1+B_2A_2)=U_1+U_2=\operatorname{Im} A_1+\operatorname{Im} A_2$. Indeed, it follows that $B_2A_2B_1A_1=0$ and $(B_2A_2)^2=B_2A_2$. Therefore, if E is the identity permutation, then we have

$$(E-B_2A_2)(B_1A_1+B_2A_2) = B_1A_1 \text{ and } B_2A_2(B_1A_1+B_2A_2) = B_2A_2.$$

4. Main theorem

Here we formulate and prove our main theorem.

Theorem 8. *If M is a finite set with $|M|\geqq 4$, H is a triply transitive permutation group on M and $f\in O_M$ is an essential operation, then either $\langle H\cup\{f\}\rangle = O_M$, or $|M|=2^n$ for some $n\geqq 2$ and $\langle H\cup\{f\}\rangle = ACL(n,2)$.*

Proof. Let M, H and f satisfy the assumptions of the theorem. If $\langle H\cup\{f\}\rangle \neq O_M$ then, by Theorem C, we have that $|M|=2^n$ for some $n\geqq 2$ and $\langle H\cup\{f\}\rangle \subseteqq ACL(n,2)$. We have to show that the latter inclusion is actually an equality. Let $\bar{H}$ denote the group of all permutations belonging to $\langle H\cup\{f\}\rangle$.

If $\bar{H}=AGL(n,2)$, then by Theorem B we have $\langle H\cup\{f\}\rangle = ACL(n,2)$. Suppose that $\bar{H}$ is a proper subgroup of $AGL(n,2)$. Then applying Theorem A we get that $n=4$, and if G denotes the subgroup of $\bar{H}$ containing all permutations of H fixing the zero vector then $G\cong \mathscr{A}_7$.

Let s be the minimum of the arities of essential operations belonging to $\langle H\cup\{f\}\rangle$ and let g be an s-ary essential operation in $\langle H\cup\{f\}\rangle$. Since H is transitive, we can suppose that $g(0,\dots,0)=0$ and thus g has the form $\sum_{i=1}^{s} x_iA_i$. We show that $s=2$. Suppose $s\geqq 3$. If for some $j\in\{1,\dots,s\}$ there is a $k\in\{1,\dots,s\}\setminus\{j\}$ such that $\mathrm{Im}\,A_j\subseteqq \mathrm{Im}\,A_k$ then $g(x_1,\dots,x_{k-1},0,x_{k+1},\dots,x_s)$ is an $(s-1)$-ary essential operation and it belongs to $\langle H\cup\{f\}\rangle$ by Lemma 3. This contradicts the assumption on s. Hence we have that $\mathrm{Im}\,A_1, \mathrm{Im}\,A_2\neq V(4,2)$, and $\mathrm{Im}\,A_1\nsubseteq \mathrm{Im}\,A_2$ and $\mathrm{Im}\,A_2\nsubseteq \mathrm{Im}\,A_1$. Then Lemma 7 yields a procedure for constructing an $(s-1)$-ary essential operation, a contradiction. Hence $s=2$, and $g(x_1,x_2)=x_1A_1+x_2A_2$.

First consider the case when $\mathrm{Im}\,A_1=V(4,2)$ (the case $\mathrm{Im}\,A_2=V(4,2)$ can be handled similarly). Then $x_1+x_2A_2=x_1A_1^{-1}A_1+x_2A_2\in\langle H\cup\{f\}\rangle$. Applying Lemmas 4, 5 and Lemma 3, one can easily show that there is a unary operation $B\in\langle H\cup\{f\}\rangle$ fixing 0 such that $\dim(\mathrm{Im}\,B)=1$, $B^2=B$ and $x_1+x_2B\in\langle H\cup\{f\}\rangle$. Choose a basis $u_1,\dots,u_4$ of $V(4,2)$ such that u_1 and u_2, u_3, u_4 are bases of $\mathrm{Im}\,B$ and $\mathrm{Ker}\,B$ respectively. Let $C\in G$ be such that u_1+u_1C, u_2, u_3, u_4 is again a basis of $V(4,2)$, and let E denote the identity permutation. Then $u_1(E+BC)=u_1+u_1C$ and $u_i(E+BC)=u_i$, $i=2,3,4$, implying that $E+BC$ is a permutation, and thus $E+BC\in G$. Hence for $E, E+BC\in G$ we have $u_iE=u_i(E+BC)$, $i=2,3,4$. Therefore by Proposition 1 it follows that $E=E+BC$ implying $BC=0$, a contradiction.

Finally consider the case when $\mathrm{Im}\,A_1, \mathrm{Im}\,A_2\neq V(4,2)$. Then Lemma 7 yields a procedure for constructing a binary operation $x_1B_1+x_2B_2\in\langle H\cup\{f\}\rangle$ such that $\mathrm{Im}(B_1+B_2)=V(4,2)$. Then $B_1+B_2\in G$ and the operation $h(x_1,x_2)=(x_1B_1+x_2B_2)(B_1+B_2)^{-1}$ is idempotent. Consider the operations $h_0(x_1,x_2)=h(x_1,x_2)$ and $h_n(x_1,x_2)=h_{n-1}(h(x_1,x_2),x_2)$ if $n\geqq 1$. It is easy to check that there is a $t\geqq 0$ such that for $h_t(x_1,x_2)=x_1C_1+x_2C_2$ we have either $C_1^2=C_1$ or

$C_1^2=0$, $C_1\neq 0$. Since h_t is idempotent, we have that $h_t(x_1, x_2)=x_1C_1+x_2(E-C_1)$. If $C_1^2=0$, then $(E-C_1)^2=E$, which shows that $\operatorname{Im}(E-C_1)=V(4,2)$, and this case has been settled.

Now suppose that $C_1^2=C_1$ and consider the operation $x_1C_1+x_2(E-C_1)$. Let $\dim(\operatorname{Im} C_1)=k$ and $\dim(\operatorname{Ker} C_1)=l$. Then clearly $1\leqq k, l$ and $k+l=4$. Choose a basis $u_1, \ldots, u_4$ of $V(4,2)$ such that $u_1, \ldots, u_k$ and $u_{k+1}, \ldots, u_4$ are bases of $\operatorname{Im} C_1$ and $\operatorname{Ker} C_1$. From now on consider the unary linear operations fixing 0 as 4×4 matrix over $GF(2)$ in the basis $u_1, \ldots, u_4$. Let D be a permutation belonging to $GL(4,2)\setminus G$. Then, by Proposition 2, there are $D_1, D_2\in G$ such that the first k columns of D and D_1 are equal, and the last l columns of D and D_2 are equal. Then it is easy to check that $D=D_1C_1+D_2(E-C_1)$ and thus $D\in G$, a contradiction. This completes the proof.

5. Application

An algebra $\langle M; F\rangle$ is said to be *homogeneous* if every permutation on M is an automorphism of $\langle M; F\rangle$. In [2] B. Csákány proved that almost all at least two element nontrivial finite algebras are functionally complete. The exceptional algebras are equivalent to one of the following six algebras:

(1) $\langle\{0,1\}; s\rangle$ where $s(x)=x+1 \pmod 2$,

(2) $\langle\{0,1\}; m\rangle$ where $m(x,y,z)=x+y+z \pmod 2$,

(3) $\langle\{0,1\}; t\rangle$ where $t(x,y,z)=x+y+z+1 \pmod 2$,

(4) $\langle\{0,1\}; d\rangle$ where $d(x,y,z)=xy+xz+yz \pmod 2$,

(5) $\langle\{0,1,2\}; l\rangle$ where $l(x,y,z)=x-y+z \pmod 3$,

(6) $\langle\{0,1\}^2; m\rangle$.

The result above was improved in [8] as follows: An at least four element nontrivial finite algebra with triply transitive automorphism group is either functionally complete or equivalent to the algebra $\langle\{0,1\}^n; m\rangle$ for some $n\geqq 2$.

A relation algebra $\langle M; R\rangle$ is said to be *relationally complete* if $[R\cup\{\{a\}|a\in M\}]=R_M$. As an analogue of Csákány's result R. Pöschel [6] proved the following: Almost all at least two element finite nontrivial homogeneous relation algebras are relationally complete. The exceptional relation algebras are equivalent to one of the following five relation algebras:

(1′) $\langle\{0,1\}; s^{\bullet}\rangle$,

(2′) $\langle\{0,1\}; m^{\bullet}\rangle$,

(3′) $\langle\{0,1\}; t^{\bullet}\rangle$,

(4′) $\langle\{0,1,2\}; l^{\bullet}\rangle$,

(5′) $\langle\{0,1\}^2; m^{\bullet}\rangle$.

Now we apply Theorem 9 to get the analogue of the result in [8] formulated above for relation algebras, which is an improvement of Pöschel's result.

Theorem 9. *An at least four element nontrivial finite relation algebra with triply transitive automorphism group is either relationally complete or equivalent to the relation algebra* $\langle \{0, 1\}^n; m^\bullet \rangle$ *for some* $n \geqq 2$.

Proof. Let $\langle M; R \rangle$ be a relation algebra satisfying the assumptions of the theorem. If $\langle M; R \rangle$ is not relationally complete, then

$$R_M \neq \big[R \cup \big\{\{a\} \mid a \in M\big\}\big] = \operatorname{Inv} \operatorname{Pol} \big(R \cup \big\{\{a\} \mid a \in M\big\}\big) =$$
$$= \operatorname{Inv} \big(\operatorname{Pol} R \cap \operatorname{Pol} \big(\big\{\{a\} \mid a \in M\big\}\big)\big) = \operatorname{Inv} (I \cap \operatorname{Pol} R)$$

where clearly $I = \operatorname{Pol} \big(\big\{\{a\} \mid a \in M\big\}\big)$ is the set of all idempotent operations in O_M. It follows that $I \cap \operatorname{Pol} R$ contains a nontrivial operation f which is evidently essential.

Now $\operatorname{Aut} \langle M; R \rangle \cup \{f\} \subseteqq \operatorname{Pol} R$ and $\operatorname{Pol} R \neq O_M$. Therefore, by Theorem 9, we have that there is an $n \geqq 2$ such that $|A| = 2^n$ and $\operatorname{Pol} R = AGL(n, 2)$. It is well-known (cf. e.g. [5]) that $\operatorname{Inv} \big(AGL(n, 2)\big) = [m^\bullet]$. Hence $\operatorname{Inv} \operatorname{Pol} R = [m^\bullet]$, which was to be proved.

References

[1] P. J. Cameron and W. M. Kantor, 2-transitive and antiflag transitive collineation groups o finite projective spaces, *J. Algebra*, **60** (1979), 384—422.

[2] B. Csákány, Homogeneous algebras are functionally complete, *Algebra Universalis*, **11** (1980), 149—158.

[3] B. Huppert, *Endliche Gruppen. I*, Springer-Verlag (Berlin—Heidelberg—New York, 1967).

[4] P. P. Pálfy, personal communication.

[5] R. Pöschel and L. A. Kalužnin, *Funktionen- und Relationenalgebren*, Deutscher Verlag der Wissenschaften (Berlin, 1979).

[6] R. Pöschel, Homogeneous relational algebras are relationally complete, in: *Finite Algebra and Multiple-Valued Logic* (Proc. Conf. Szeged, 1979), Colloq. Math. Soc. J. Bolyai, vol. 28, North-Holland (Amsterdam, 1981); pp. 587—601.

[7] P. Schofield, Complete subsets of mappings over a finite domain, *Proc. Camb. Phil. Soc.*, **62** (1966), 597—611.

[8] L. Szabó and Á. Szendrei, Almost all algebras with triply transitive automorphism group are functionally complete, *Acta Sci. Math.*, **41** (1979), 391—402.

[9] L. Szabó and Á. Szendrei, Słupecki-type criteria for quasilinear functions over a finite dimensional vector space, *Elektron. Informationsverarb. und Kybernet.*, **17** (1981), 601—611.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math., 51 (1987), 229—249

A generalization of McAlister's P-theorem for E-unitary regular semigroups

MÁRIA B. SZENDREI

To the memory of András Huhn

A regular semigroup is called E-unitary if its set of idempotents is a unitary subset. One can easily show that E-unitary regular semigroups are necessarily orthodox.

In 1974 McALISTER [5], [6] proved that every inverse semigroup is an idempotent separating homomorphic image of an E-unitary inverse semigroup and described E-unitary inverse semigroups by means of groups, partially ordered sets and semilattices. This structure theorem is referred to as the "P-theorem". By making use of McAlister's P-theorem O'CARROLL [8] proved that every E-unitary inverse semigroup can be embedded into a semidirect product of a semilattice by a group.

These results have opened up new perspectives not only in the theory of inverse semigroups but in the theory of regular semigroups. McAlister's first result was generalized for orthodox semigroups independently by TAKIZAWA [15] and the author [10]. TAKIZAWA [14] generalized the P-theorem, too, but only for E-unitary $\mathscr{R}$-unipotent semigroups. This structure theorem was applied in [12] to prove the analogue of O'Carroll's embedding theorem for E-unitary $\mathscr{R}$-unipotent semigroups.

The aim of this paper is to present a generalization of the P-theorem for E-unitary regular semigroups. It has to be pointed out in advance that our main result which is proved in Sections 2 and 3 cannot be considered as a structure theorem in the sense that E-unitary regular semigroups are constructed in it from "simpler" objects. Indeed, it is doubtful that strictly combinatorial semigroups which play an important role in the construction are "simpler" than E-unitary regular semigroups. However, the strictly combinatorial partial semigroup introduced in Section 2 is applied in a forthcoming paper [13] to prove that every E-unitary regular semigroup with regular band of idempotents can be embedded into a semidirect product of a band by a group.

Received July 5, 1985.

Margolis and Pin [4] generalized McAlister's P-theorem in another direction, namely for E-unitary not necessarily regular semigroups with commuting idempotents. It turns out that in the special case of E-unitary regular semigroups with commuting idempotents, that is, in the case of E-unitary inverse semigroups the main theorem of [4] asserts almost the same result as a part of our main theorem. In Section 4 we deduce a characterization of E-unitary regular semigroups which is similar to that formulated in the main theorem of [4].

1. Preliminaries

Let S be a semigroup. The set of idempotents in S is denoted by E_S and the set of inverses of an element s in S by $V_S(s)$. For the least group congruence on S we use the notation σ_S and the factor semigroup S/σ_S will be denoted by G_S. If it causes no confusion we omit S from E_S, $V_S(s)$ and σ_S.

A regular semigroup S is called *E-unitary* if E is a unitary subset in S. It is easy to see that E-unitary regular semigroups are necessarily orthodox.

Result 1.1 (Howie and Lallement [3] and Saitô [9]). *For a regular semigroup S, the following conditions are equivalent:*

(i) *S is E-unitary,*

(ii) *E is a left unitary subset in S,*

(iii) *E is a right unitary subset in S,*

(iv) *E constitutes a σ-class.*

Let $\varphi\colon S\to T$ be a homomorphism where S and T are regular semigroups. We denote by $\ker\varphi$ the congruence on S induced by φ and by $\operatorname{Ker}\varphi$ the union of idempotent $\ker\varphi$-classes. If $\varkappa$ is a congruence on S then instead of $\operatorname{Ker}\varkappa^\natural$ we simply write $\operatorname{Ker}\varkappa$.

Now let S be an orthodox semigroup with 0. Assume that S is categorical at 0. It is obvious that the least inverse semigroup congruence γ on S is 0-restricted and S/γ is also categorical at 0. Hence it follows by Theorem 7.66 [1] that there exists a least 0-restricted congruence β on S/γ such that $(S/\gamma)/\beta$ is a primitive inverse semigroup. It is easily seen that $\ker\gamma^\natural\beta^\natural$ is the least 0-restricted primitive inverse semigroup congruence on S which will be denoted by ϱ_S or, simply, by ϱ.

Proposition 1.2. *Let $S=S^0$ be an orthodox semigroup which is categorical at 0. Then the following conditions are equivalent for $s,t\in S$:*

(i) *$s,t\neq 0$ and $s\varrho t$;*

(ii) *$set'\in E\setminus 0$ for some $e\in E$ and $t'\in V(t)$;*

(iii) *$s'et\in E\setminus 0$ for some $e\in E$ and $s'\in V(s)$;*

(iv) *$se=ft\neq 0$ for some $e,f\in E\setminus 0$;*

(v) *$EsE\cap EtE\neq\{0\}$.*

Proof. One can prove the equivalence of conditions (ii)—(v) in the same fashion as the equivalence of conditions (2), (3), (6) and (8) in Lemma 1.3 [14]. One needs only to investigate whether products are 0 or not. Let us see, for example, the proof of the implication (v)$\Rightarrow$(ii). Suppose that $esf=gth\neq 0$ for some $e, f, g, h\in E$, and let $s'\in V(s)$, $t'\in V(t)$. Then $(s'es)f\in E$ and $s((s'es)f)t'=(ss')g(tht')\in E$. If it were 0 then

$$0 = e(ss')g(tht')th = e(ss')gt(t'th)^2 = e(ss')gth =$$

$$= e(ss')esf = (ess')^2 sf = (ess')sf = esf$$

would follow, a contradiction. Thus $s((s'es)f)t'\in E\setminus 0$.

Similarly to the proof of Lemma 1.3 [14], one can check that the relation $\varkappa$ consisting of the pair (0, 0) and the pairs (s, t) satisfying (ii)—(v) is a 0-restricted congruence on S. We intend to show that $\varkappa=\varrho$. First observe that $S/\varkappa$ is a primitive inverse semigroup. Indeed, if $e, f\in E$ with $ef\neq 0$ then $ef=eef=eff\in EeE\cap EfE$ and hence $e\varkappa f$. Now let τ be any 0-restricted primitive inverse semigroup congruence on S and let $e, f\in E$ with $se=ft\neq 0$. Then $s\tau\cdot e\tau=f\tau\cdot t\tau\neq 0$ in the primitive inverse semigroup S/τ. Hence we infer that $(s\tau)^{-1}\cdot s\tau=e\tau=f\tau=t\tau\cdot(t\tau)^{-1}$ which implies $s\tau=s\tau\cdot e\tau=f\tau\cdot t\tau=t\tau$. Thus $\varkappa\subseteq\tau$, completing the proof of the fact that $\varkappa=\varrho$.

A regular semigroup S with 0 is called *$E\setminus 0$-unitary* if $E\setminus 0$ is a unitary subset in S. Let S be an $E\setminus 0$-unitary regular semigroup with 0. If $e\in E\setminus 0$ and $e'\in V(e)$ then $ee'\in E\setminus 0$. Since $E\setminus 0$ is a left unitary subset in S we deduce that $e'\in E\setminus 0$. Thus S is orthodox.

Proposition 1.3. *Every $E\setminus 0$-unitary regular semigroup with 0 is orthodox.*

Thus there exists a least 0-restricted primitive inverse semigroup congruence on every $E\setminus 0$-unitary regular semigroup being categorical at 0. The analogue of Result 1.1 holds:

Proposition 1.4: *Let $S=S^0$ be an orthodox semigroup which is categorical at 0. Then the following conditions are equivalent:*

(i) *S is $E\setminus 0$-unitary;*

(ii) *$E\setminus 0$ is a left unitary subset in S;*

(iii) *$E\setminus 0$ is a right unitary subset in S;*

(iv) $\operatorname{Ker}\varrho = E$.

Proof. The equivalence of conditions (ii) and (iv) is easily verified by making use of the equivalence of (i) and (iii), (iv) in Proposition 1.2. The equivalence (iii)$\Leftrightarrow$(iv) follows by symmetry, and (i) is equivalent to (iii) and (ii) by definition.

For an $E\setminus 0$-unitary regular semigroup S which is categorical at 0, the congruence ϱ can be described as follows:

Proposition 1.5. *Let $S=S^0$ be an $E\setminus 0$-unitary regular semigroup which is categorical at 0. Then*

$$\varrho = \{(s, t)\colon st' \in E\setminus 0 \textit{ for some } t' \in V(t)\} \cup \{(0, 0)\}.$$

Proof. Denote the relation on the right hand side of the equality by $\varkappa$. It is clear by Proposition 1.2 that $\varkappa \subseteq \varrho$. Suppose now that $s, t \neq 0$ and $s\varrho t$ in S. Then there exist $e, f \in E$ with $es = tf \neq 0$. This implies $est' = tft' \in E\setminus 0$. Since S is $E\setminus 0$-unitary we obtain that $st' \in E\setminus 0$, that is, $s\varkappa t$. Thus the reverse inclusion $\varrho \subseteq \varkappa$ also holds.

In Sections 2, 3 and 4 we will need the following facts:

Lemma 1.6. *Let $S=S^0$ be an orthodox semigroup which is categorical at 0. If there exists a 0-restricted homomorphism φ of S onto a primitive inverse semigroup such that* $\operatorname{Ker} \varphi \subseteq E$ *then S is $E\setminus 0$-unitary and* $\ker \varphi = \varrho$.

Proof. Since $\ker \varphi$ is a 0-restricted primitive inverse semigroup congruence we have $\varrho \subseteq \ker \varphi$. Therefore $\operatorname{Ker} \varrho \subseteq \operatorname{Ker} \varphi \subseteq E$. However, $E \subseteq \operatorname{Ker} \varrho$ trivially holds whence we infer $\operatorname{Ker} \varrho = \operatorname{Ker} \varphi = E$. Then, by Proposition 1.4, it follows that S is $E\setminus 0$-unitary. Let $s, t \in S\setminus 0$ be such that $s\varphi = t\varphi$ and let $t' \in V(t)$. Then $(st')\varphi = s\varphi \cdot (t\varphi)^{-1} = t\varphi \cdot (t\varphi)^{-1} \in E_{S\varphi}$ which implies $st' \in E_S = E$. Thus, by Proposition 1.5, we have $s\varrho t$, completing the proof of the inclusion $\ker \varphi \subseteq \varrho$.

In order to simplify the notations later on, we will denote by $B(I)$ the $\mathscr{H}$-trivial Brandt semigroup $(I \times I) \cup 0$ with multiplication

$$[i, j][k, l] = \begin{cases} [i, l] & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$$

and

$$0[i, j] = [i, j]0 = 0 \cdot 0 = 0.$$

It is well known that every $\mathscr{H}$-trivial Brandt semigroup is isomorphic to $B(I)$ for some set I.

Lemma 1.7. *Let $S=S^0$ be an $E\setminus 0$-unitary regular semigroup which is categorical at 0 and for which S/ϱ is an $\mathscr{H}$-trivial Brandt semigroup. Then the only 0-restricted primitive inverse semigroup congruence on S is ϱ.*

Proof. A 0-restricted primitive inverse semigroup congruence properly containing ϱ cannot exist as $\mathscr{H}$-trivial Brandt semigroups are congruence-free.

If S is a semigroup with 0 then the partial groupoid obtained from S by eliminating 0 and letting products be undefined if they are equal to 0 in S will be denoted by $\hat{S}$.

Given a partial groupoid $(X;\ \cdot)$, let us adjoin a new symbol $0(\notin X)$ to X and extend the multiplication to $X\cup 0$ in such a way that $x\cdot 0=0\cdot x=0\cdot 0=0$ for every $x\in X$ and $x\cdot y=0$ provided $x, y\in X$ and $x\cdot y$ is not defined in X. The groupoid obtained in this fashion is denoted by $\check{X}$. If $\check{X}$ is a semigroup then we term X a *partial semigroup*.

The basic concepts of semigroup theory such as left, right ideals, Green's relations, inverse of an element, regularity, automorphisms can be defined in a partial semigroup X in the same way as in $\check{X}$. For example, a non-empty subset $R\subseteq X$ is said to be a *right ideal* in X if $\{r\cdot x\colon\ r\in R,\ x\in X$ and $r\cdot x$ is defined$\}\subseteq R$. Clearly, R is a right ideal in X if and only if $R\cup 0$ is a non-trivial right ideal in $\check{X}$. One can easily see that, for example, the set of all idempotent elements in X is $E_{\check{X}}\setminus 0$, Green's relation $\mathcal{R}$ on X is just the restriction of the $\mathcal{R}$-relation of $\check{X}$ to X, the set of inverses of an element x in X is equal to $V_{\check{X}}(x)$ and $\alpha\colon X\to X$ is an automorphism of X if and only if $\check{\alpha}\colon \check{X}\to\check{X}$ defined by $0\check{\alpha}=0$ and $x\check{\alpha}=x\alpha$ $(x\in X)$ is an automorphism. Therefore it is not ambiguous to write $\mathcal{R}$ or $V(x)$ without indicating whether they are considered on X or on $\check{X}$. If we want to emphasize that the set of inverses is considered in X then we write $V_X(x)$. Moreover, we will use the notation E_Y for the set of all idempotent elements in a subset Y of X and $V_X(Y)$ or, simply, $V(Y)$ for $\bigcup\{V_X(a)\colon\ a\in Y\}$.

Let G be a group and S a full or partial semigroup. We say that G *acts on* S if a homomorphism $\varphi\colon\ G\to(\operatorname{Aut} S)^d$ is given where $(\operatorname{Aut} S)^d$ is the dual of the automorphism group of S. For every $s\in S$ and $g\in G$, we denote $s(g\varphi)$ by gs.

Let G be a group and S a semigroup with 0 on which G acts. Define a multiplication on the set $((S\setminus 0)\times G)\cup 0$ by

$$(s,\ g)(t,\ h)=\begin{cases}(s\cdot gt,\ gh) & \text{if}\quad s\cdot gt\neq 0\\ 0 & \text{otherwise}\end{cases}$$

and

$$0\cdot(s,\ g)=(s,\ g)\cdot 0=0\cdot 0=0$$

for every $s,\ t\in S\setminus 0$ and $g,\ h\in G$. It is not difficult to check that this multiplication is associative. The semigroup obtained in this way is called the *0-semidirect product of S by G* and is denoted by $S*_0G$.

If G is a group acting on a semigroup S without 0 then $(S^0*_0G)\setminus 0$ is a semigroup termed the *semidirect product of S by G* and is denoted by $S*G$.

Let X be a partial semigroup and G a group acting on X. Let $\varphi\colon\ G\to(\operatorname{Aut} X)^d$ be the homomorphism defining this action. Then $\check{\varphi}\colon\ G\to(\operatorname{Aut}\check{X})^d$, $g\check{\varphi}=\widecheck{g\varphi}$ is a homomorphism. Since $x(g\check{\varphi})=x(g\varphi)$ for every $x\in X$ and $g\in G$, it is not confusing to denote $x(g\check{\varphi})$ also by gx. By the *semidirect product* $X*G$ we mean the partial semigroup $\widehat{\check{X}*_0G}$.

2. On E-unitary regular semigroups

By McAlister's P-theorem [6], every E-unitary inverse semigroup S is isomorphic to a P-semigroup $P(G, \mathscr{X}, \mathscr{Y})$ where G is a group, $\mathscr{X}$ is a partially ordered set on which G acts by order automorphisms, $\mathscr{Y}$ is an order ideal in $\mathscr{X}$ such that $\mathscr{Y}$ is a lower semilattice and $P(G, \mathscr{X}, \mathscr{Y})$ is, actually, a well-determined subsemigroup in the semidirect product of the "partial semilattice" $\mathscr{X}$ by G. Takizawa ([14]; cf. also [11]) generalized this result by proving that every E-unitary $\mathscr{R}$-unipotent semigroup S is isomorphic to a so-called PL-semigroup constructed in a similar way as a P-semigroup by means of a group, an $\mathscr{R}$-trivial "partial idempotent semigroup" $\mathscr{X}$ on which G acts and by means of a subband $\mathscr{Y}$ of $\mathscr{X}$ forming an order ideal in $\mathscr{X}$ with respect to the natural partial order $\leqq_{\mathscr{X}}$. In both cases the triple $(G, \mathscr{X}, \mathscr{Y})$ can be chosen in such a way that G is isomorphic to G_S and $\mathscr{Y}$ to E_S.

The proofs of McAlister's and Takizawa's results are based on the observation that $\mathscr{R}\cap\sigma=\iota$ (ι is the identity relation) holds in an E-unitary inverse [$\mathscr{R}$-unipotent] semigroup (cf. [6] and [14]). Hence the elements of an E-unitary inverse [$\mathscr{R}$-unipotent] semigroup S can be coordinatized with pairs from $E_S\times G_S$.

When we intend to generalize these results for E-unitary regular semigroups the difficulty lies in the fact that, in an arbitrary E-unitary regular semigroup S, we have no such natural coordinatization of elements as in the case of E-unitary inverse [$\mathscr{R}$-unipotent] semigroups. The analogue of that coordinatization would be the injection $S\to E_S/\mathscr{R}\times G_S\times E_S/\mathscr{L}$ defined by $s\mapsto(R_{ss'}, s\sigma, L_{s's})$ where $s'\in V(s)$. However, it seems very complicated to determine in an abstract way which triples are coordinates of an element, how the coordinates are multiplied and what conditions they have to satisfy in order that the groupoid defined in this way be an E-unitary regular semigroup. Therefore we looked for another way of characterizing E-unitary regular semigroups. We cannot expect to obtain a construction analogous to P-semigroups which produced all E-unitary regular semigroups up to isomorphisms and in which $\mathscr{Y}$ were isomorphic to E_S. In finding a generalization of the P-theorem for E-unitary regular semigroups, we tried to preserve the other main feature of McAlister's and Takizawa's results, namely, we wanted to obtain an E-unitary regular semigroup as a well-determined subsemigroup of a semidirect product of a certain partial groupoid by a group. We imitate the proof of the P-theorem due to Munn [7] and that of Theorem 3.1 in [14]. The new idea in our case is that the partial groupoid $\mathscr{X}$ is defined on $S\times G_S$ instead of $E_S\times G_S$.

Let S be an E-unitary regular semigroup. Define a partial groupoid $\mathscr{X}=(S\times G_S;\circ)$ as follows:

$$(1)\qquad \begin{gathered}(s, g)\circ(t, h) \text{ is defined if and only if } s\sigma = g^{-1}h,\\ \text{and in this case } (s, g)\circ(t, h) = (st, g).\end{gathered}$$

Put $\mathscr{Y}=\{(s, 1)\colon s\in S\}$.

In the sequel we prove several properties of the triple $(G_S, \mathscr{X}, \mathscr{Y})$.

(I) $\check{\mathscr{X}}$ *is an orthodox semigroup which is categorical at* 0. *Moreover,* $E_{\mathscr{X}}=E_S\times G_S$ *and, for every* $(s, g)\in\mathscr{X}$, *we have* $V_{\mathscr{X}}((s, g))=\{(s', g\cdot s\sigma)\colon s'\in V_S(s)\}$.

Proof. Let $(s, g), (t, h), (u, k)\in\mathscr{X}$. It is clear by (1) that we have $(s, g)\circ(t, h)=0$ in $\check{\mathscr{X}}$ if and only if $s\sigma\neq g^{-1}h$. Suppose first that $(s, g)\circ(t, h)\neq 0$ and $(t, h)\circ(u, k)\neq 0$. Then $s\sigma=g^{-1}h$ and $t\sigma=h^{-1}k$ which imply that $(st)\sigma=g^{-1}k$. Hence it follows by (1) that $((s, g)\circ(t, h))\circ(u, k)=(st, g)\circ(u, k)=(stu, g)=(s, g)\circ(tu, h)=(s, g)\circ((t, h)\circ(u, k))\neq 0$. If $(s, g)\circ(t, h)=0$ and $(t, h)\circ(u, k)\neq 0$, then $s\sigma\neq g^{-1}h$ which implies by (1) that $(s, g)\circ((t, h)\circ(u, k))=(s, g)\circ(tu, h)=0$. If $(s, g)\circ(t, h)\neq 0$ and $(t, h)\circ(u, k)=0$, then $s\sigma=g^{-1}h$ and $t\sigma\neq h^{-1}k$ whence we infer that $(st)\sigma\neq g^{-1}k$. Therefore $((s, g)\circ(t, h))\circ(u, k)=(st, g)\circ(u, k)=0$. Thus we have shown that $\check{\mathscr{X}}$ is a semigroup which is categorical at 0.

Let $(s, g)\in\mathscr{X}$. Now we determine $V_{\mathscr{X}}((s, g))$. Making use of the fact that $s'\sigma=(s\sigma)^{-1}$ for each $s'\in V_S(s)$, one can easily check that $(s', g\cdot s\sigma)\in V_{\mathscr{X}}((s, g))$ for every $s'\in V_S(s)$. If $(t, h)\in V_{\mathscr{X}}((s, g))$ then (1) implies $t\in V_S(s)$ and, since $(s, g)\circ(t, h)\neq 0$, we have $s\sigma=g^{-1}h$. So it is verified that $V_{\mathscr{X}}((s, g))$ consists of those elements indicated in the assertion. In particular, we obtain that $\check{\mathscr{X}}$ is regular.

It remains to determine $E_{\mathscr{X}}$. It is obvious that $(e, g)\in E_{\mathscr{X}}$ for any $e\in E_S$ and $g\in G_S$. Assume that $(e, g)\in E_{\mathscr{X}}$. Then $(e, g)\circ(e, g)=(e, g)$, that is, $e\sigma=g^{-1}g=1$ and $e^2=e$. Clearly, $E_{\check{\mathscr{X}}}$ is a band because, for every $(e, g), (f, h)\in E_{\mathscr{X}}$, we have

$$(2)\qquad (e, g)\circ(f, h) = \begin{cases}(ef, g) & \text{if } g=h\\ 0 & \text{otherwise.}\end{cases}$$

This implies that $\check{\mathscr{X}}$ is orthodox. The proof is complete.

(II) *The mapping* $\varphi\colon \check{\mathscr{X}}\to B(G_S)$ *defined by* $(s, g)\varphi=[g, g\cdot s\sigma]$ *and* $0\varphi=0$ *is a surjective* 0*-restricted homomorphism with* $\operatorname{Ker}\varphi\subseteq E_{\check{\mathscr{X}}}$. *Consequently,* $\check{\mathscr{X}}$ *is* $E\setminus 0$*-unitary,* $\ker\varphi=\varrho$, *the least* 0*-restricted primitive inverse semigroup congruence on* $\check{\mathscr{X}}$ *and* $\check{\mathscr{X}}/\varrho$ *is an* $\mathscr{H}$*-trivial Brandt semigroup.*

Proof. Let $(s, g), (t, h)\in\mathscr{X}$. If $s\sigma=g^{-1}h$, that is, $g\cdot s\sigma=h$ then

$$((s, g)\circ(t, h))\varphi = (st, g)\varphi = [g, g\cdot(st)\sigma] =$$
$$= [g, h\cdot t\sigma] = [g, g\cdot s\sigma]\cdot[h, h\cdot t\sigma] = (s, g)\varphi\cdot(t, h)\varphi.$$

If $s\sigma\neq g^{-1}h$, that is, $g\cdot s\sigma\neq h$ then

$$((s, g)\circ(t, h))\varphi = 0\varphi = 0 = [g, g\cdot s\sigma][h, h\cdot t\sigma] = (s, g)\varphi\cdot(t, h)\varphi.$$

Thus φ is a 0-restricted homomorphism. It is surjective because $G_S=S/\sigma$. Since S is E-unitary, $s\sigma=1$ implies $s\in E_S$. Therefore, by (I), $\operatorname{Ker}\varphi\subseteq E_{\check{\mathscr{X}}}$. By Lemma

1.6, this ensures that $\check{\mathscr{X}}$ is $E\setminus 0$-unitary and $\ker\varphi$ is the least 0-restricted primitive inverse semigroup congruence ϱ.

(III) *$\mathscr{Y}$ is a maximal right ideal in $\mathscr{X}$ with the property that $E_{\mathscr{Y}}$ is a subband in $\mathscr{Y}$.*

Proof. It is straightforward by (1) and (2) that $\mathscr{Y}$ is a right ideal in $\mathscr{X}$ and $E_{\mathscr{Y}}$ is a subband in $\mathscr{Y}$. Suppose now that $\mathscr{Y}_1$ is a right ideal in $\mathscr{X}$ such that $E_{\mathscr{Y}_1}$ is a subband in $\mathscr{Y}_1$ and $\mathscr{Y}\subseteqq\mathscr{Y}_1$. Then (2) implies $E_{\mathscr{Y}}=E_{\mathscr{Y}_1}$. Let $(s,g)\in\mathscr{Y}_1$. Since $\mathscr{Y}_1$ is a right ideal in $\mathscr{X}$ we infer by (I) that $(ss',g)=(s,g)\circ(s',g\cdot s\sigma)\in\mathscr{Y}_1\cap E_{\mathscr{X}}=$ $=E_{\mathscr{Y}_1}=E_{\mathscr{Y}}$ for every $s'\in V_S(s)$. Thus $g=1$ and $(s,g)\in\mathscr{Y}$ proving that $\mathscr{Y}_1\subseteqq\mathscr{Y}$. The proof is complete.

Let us define an action of G_S on $\mathscr{X}$ as follows: for every (s,g) and $h\in G_S$ let $h(s,g)=(s,hg)$.

(IV) *G_S acts on $\mathscr{X}$ such that $G_S\mathscr{Y}=\mathscr{X}$ and, for every $g\in G_S$, there exists $a\in\mathscr{Y}$ with $ga\in V_{\mathscr{X}}(\mathscr{Y})$.*

Proof. By (1), one can immediately check that, for every $h\in G_S$, the mapping $\tilde{h}$: $\mathscr{X}\to\mathscr{X}$ defined by $(s,g)\tilde{h}=(s,hg)$ is an automorphism and $\widetilde{kh}=\tilde{h}\tilde{k}$ for every $h,k\in G_S$. The equality $G_S\mathscr{Y}=\mathscr{X}$ is a trivial consequence of the definition of the action. In order to verify the last assertion, observe that, by (I), we have $V_{\mathscr{X}}(\mathscr{Y})=$ $=\{(s',s\sigma)\colon s\in S,\ s'\in V_S(s)\}=\{(t,h)\in\mathscr{X}\colon t\sigma=h^{-1}\}$. Since $G_S=S/\sigma$, for any $g\in G_S$, there exists $s\in S$ with $s\sigma=g^{-1}$. For such an s we have $g(s,1)=(s,g)\in V_{\mathscr{X}}(\mathscr{Y})$.

As an easy consequence of the equality obtained here for $V_{\mathscr{X}}(\mathscr{Y})$ we deduce

(V) *For every $(s,1)\in\mathscr{Y}$ and $g\in G_S$, we have $g^{-1}(s,1)\in V_{\mathscr{X}}(\mathscr{Y})$ if and only if $s\sigma=g$.*

(VI) *The mapping ε: $S\to\mathscr{X}*G_S$ defined by $s\varepsilon=((s,1),s\sigma)$ is an embedding of S into $\mathscr{X}*G_S$. In particular, S is isomorphic to the subsemigroup $\{(a,g)\in\mathscr{Y}\times G_S\colon g^{-1}a\in V_{\mathscr{X}}(\mathscr{Y})\}$.*

Proof. The mapping ε is clearly injective and, by (V), its range is $\{(a,g)\in$ $\in\mathscr{Y}\times G_S\colon g^{-1}a\in V_{\mathscr{X}}(\mathscr{Y})\}$. All we have to check is that ε is a homomorphism. Let $s,t\in S$. Then, by (1), we have

$$s\varepsilon\cdot t\varepsilon=((s,1),s\sigma)((t,1),t\sigma)=((s,1)\circ s\sigma(t,1),s\sigma\cdot t\sigma)=$$
$$=((s,1)\circ(t,s\sigma),(st)\sigma)=((st,1),(st)\sigma)=(st)\varepsilon$$

which completes the proof.

Statement (VI) shows that we succeeded in finding a partial semigroup $\mathscr{X}$ on which G_S acts such that S is isomorphic to a well-determined subsemigroup of $\mathscr{X}*G_S$.

3. PO-semigroups and 0-semidirect products of strictly combinatorial semigroups by groups

In this section we introduce the concept of a PO-triple and a PO-semigroup so as it is inspired by the results of the preceding section and give a description of E-unitary regular semigroups by means of PO-semigroups and by means of 0-semidirect products of strictly combinatorial semigroups by groups.

A regular semigroup S with 0 is called *strictly combinatorial* if (i) S is categorical at 0, (ii) S is $E\setminus 0$-unitary and (iii) S/ϱ is an $\mathscr{H}$-trivial Brandt semigroup.

A partial semigroup X is termed *strictly combinatorial* if $\check{X}$ is a strictly combinatorial semigroup.

An $\mathscr{H}$-trivial semigroup is sometimes called combinatorial. In order to justify the terminology just introduced we show that a strictly combinatorial semigroup is necessarily $\mathscr{H}$-trivial. Let S be a strictly combinatorial semigroup and s an element in a non-zero subgroup of S. Then there exists an inverse s' of s in this subgroup and thus $ss'=s's\neq 0$. Hence we have $(s\varrho)(s\varrho)^{-1}=(s\varrho)^{-1}(s\varrho)\neq 0$ in the factor semigroup S/ϱ which is an $\mathscr{H}$-trivial Brandt semigroup. This implies that $s\varrho$ is idempotent and thus $s\in \operatorname{Ker} \varrho\setminus 0$. Since S is $E\setminus 0$-unitary, we infer by Proposition 1.4 that s is idempotent. Thus we verified that each subgroup in S is trivial which implies that S is $\mathscr{H}$-trivial.

Now we define the notions which will play the role of the McAlister triple and the P-semigroup.

Let G be a group, $(\mathscr{X}; \circ)$ a strictly combinatorial partial semigroup and $\mathscr{Y}$ a subset in $\mathscr{X}$. Suppose that

(PO1) $\mathscr{Y}$ is a right ideal in $(\mathscr{X}; \circ)$ and $E_{\mathscr{Y}}$ is a subband in $\mathscr{Y}$;

(PO2) G acts on $(\mathscr{X}; \circ)$;

(PO3) $G\mathscr{Y} = \mathscr{X}$;

(PO4) for every $g\in G$, there exists $a\in\mathscr{Y}$ with $ga\in V(\mathscr{Y})$.

The triple $(G, \mathscr{X}, \mathscr{Y})$ satisfying the above conditions is called a *PO-triple*. If

(M) $\mathscr{Y}$ is a maximal right ideal in $(\mathscr{X}; \circ)$ with the property that $E_{\mathscr{Y}}$ is a subband in $\mathscr{Y}$, then $(G, \mathscr{X}, \mathscr{Y})$ is termed a *POM-triple*.

Given a PO-triple $(G, \mathscr{X}, \mathscr{Y})$, we define a multiplication on the set

$$PO(G, \mathscr{X}, \mathscr{Y}) = \{(a, g)\in\mathscr{Y}\times G\colon\ g^{-1}a\in V(\mathscr{Y})\}$$

by

(3) $$(a, g)(b, h) = (a\circ gb, gh).$$

Property (PO4) ensures that the image of $PO(G, \mathscr{X}, \mathscr{Y})$ under the second projection is just G. The following lemma shows that the image of $PO(G, \mathscr{X}, \mathscr{Y})$ under the first projection is $\mathscr{Y}$.

Lemma 3.1 *Let $(G, \mathscr{X}, \mathscr{Y})$ be a PO-triple. Then, for every $a\in\mathscr{Y}$, there exists $g\in G$ with $ga\in V(\mathscr{Y})$.*

Proof. If $a\in\mathscr{Y}$ and $a'\in V(a)$, then, by (PO3), we have $a'=hb$ for some $h\in G$ and $b\in\mathscr{Y}$. Thus, by (PO2), $h^{-1}a\in V(h^{-1}a')=V(b)\subseteq V(\mathscr{Y})$.

Proposition 3.2. *Let $(G, \mathscr{X}, \mathscr{Y})$ be a PO-triple.*

(i) *$PO(G, \mathscr{X}, \mathscr{Y})$ is an E-unitary regular semigroup and $E_{PO(G,\mathscr{X},\mathscr{Y})}=\{(e, 1): e\in E_{\mathscr{Y}}\}$ is isomorphic to $E_{\mathscr{Y}}$.*

Moreover, for any $(a, g), (b, h)\in PO(G, \mathscr{X}, \mathscr{Y})$, we have

(ii) $V_{PO(G,\mathscr{X},\mathscr{Y})}((a, g))=\{(c, g^{-1}): c\in\mathscr{Y}\cap V(g^{-1}a)\}$;

(iii) $(a, g)\mathscr{R}(b, h)$ *if and only if* $a\mathscr{R}b$;

(iv) $(a, g)\mathscr{L}(b, h)$ *if and only if* $g^{-1}a\mathscr{L}h^{-1}b$;

(v) $(a, g)\gamma(b, h)$ *if and only if* $g=h$ *and* $V(g^{-1}a)\cap V(h^{-1}b)\cap\mathscr{Y}\neq\square$;

(vi) $(a, g)\sigma(b, h)$ *if and only if* $g=h$.

(vii) $PO(G, \mathscr{X}, \mathscr{Y})/\sigma$ *is isomorphic to* G.

The E-unitary regular semigroup $PO(G, \mathscr{X}, \mathscr{Y})$ is called the *PO-semigroup determined by the PO-triple* $(G, \mathscr{X}, \mathscr{Y})$ or, simply, a *PO-semigroup*.

Proof. For brevity, denote $PO(G, \mathscr{X}, \mathscr{Y})$ by S.

(i) First of all, we have to show that S is closed under the multiplication defined by (3). Let $(a, g), (b, h)\in S$. Then $g^{-1}a\in V(a^+)$ and $h^{-1}b\in V(b^+)$ for some $a^+, b^+\in\mathscr{Y}$. Since $\mathscr{Y}$ is a right ideal in $\mathscr{X}$ this implies by (PO2) that $a^+\circ g^{-1}a$ and $b\circ hb^+$ belong to $E_{\mathscr{Y}}$. As $E_{\mathscr{Y}}$ is a subband in $\mathscr{Y}$ by (PO1), the product $(a^+\circ g^{-1}a)\circ(b\circ hb^+)$ is defined and thus $g^{-1}a\circ(a^+\circ g^{-1}a)\circ(b\circ hb^+)\circ b=g^{-1}a\circ b$ is also defined. From this it follows by (PO2) that $a\circ gb$ is defined and, since $\mathscr{Y}$ is a right ideal in $\mathscr{X}$, it belongs to $\mathscr{Y}$. Moreover, we obtain that $(gh)^{-1}(a\circ gb)=h^{-1}(g^{-1}a)\circ h^{-1}b$ is also defined in $\mathscr{X}$, that is, it is not equal to 0 in $\check{\mathscr{X}}$. Since the strictly combinatorial semigroup $\check{\mathscr{X}}$ is orthodox by Proposition 1.3, we infer that $h^{-1}(g^{-1}a)\circ h^{-1}b\in V(b^+\circ h^{-1}a^+)$. Hence $b^+\circ h^{-1}a^+\neq 0$, that is, the product $b^+\circ h^{-1}a^+$ is defined in $\mathscr{X}$. Since $\mathscr{Y}$ is a right ideal in $\mathscr{X}$ and $b^+\in\mathscr{Y}$, we have $b^+\circ h^{-1}a^+\in\mathscr{Y}$. Thus $(gh)^{-1}(a\circ gb)\in V(\mathscr{Y})$, completing the proof of the fact that S is closed under multiplication (3).

A straightforward calculation shows that the multiplication defined by (3) is associative. Now we turn to proving the regularity of S. Observe that it suffices to verify (ii). For, if $(a, g)\in S$ then $g^{-1}a\in V(\mathscr{Y})$. Therefore there exists $b\in\mathscr{Y}\cap V(g^{-1}a)$ and hence $gb\in V(a)\subseteq V(\mathscr{Y})$. Thus $(b, g^{-1})\in S$ is an inverse of (a, g). The element (b, h) is an inverse of (a, g) if and only if $(a, g)=(a, g)(b, h)(a, g)=(a\circ gb\circ gha, ghg)$ and $(b, h)=(b, h)(a, g)(b, h)=(b\circ ha\circ hgb, hgh)$, that is, if and only if $h=g^{-1}$ and $a\circ gb\circ a=a, b\circ g^{-1}a\circ b=b$. The latter equalities are equivalent by (PO2) to the condition that $b\in V(g^{-1}a)$.

It is easy to see that $E_S=\{(e, 1)\colon e\in E_{\mathscr{Y}}\}$ which is a band with respect to the multiplication defined in (3) and E_S is isomorphic to $E_{\mathscr{Y}}$.

Now we prove that the homomorphism $\varphi\colon S\to G$, $(a,g)\varphi=g$ is onto and $\operatorname{Ker}\varphi\subseteqq E_S$. This implies that $\ker\varphi=\sigma$ and thus, by Result 1.1, S is E-unitary and (vi), (vii) hold. Property (PO4) of the PO-triple $(G, \mathscr{X}, \mathscr{Y})$ ensures that φ is onto. Assume that $(a,g)\in\operatorname{Ker}\varphi$. Then $g=1$ and $(a,1)^2\in S$ whence $a\circ a\in\mathscr{Y}$. Since $\check{\mathscr{X}}/\varrho$ is and $\mathscr{H}$-trivial Brandt semigroup and in such a Brandt semigroup a square of an element x is non-zero if and only if x is idempotent we infer that $a\varrho$ is a non-zero idempotent in $\check{\mathscr{X}}/\varrho$. However, $\operatorname{Ker}\varrho=E_{\check{\mathscr{X}}}$ by Proposition 1.4. This implies $a\in E_{\mathscr{X}}$. Thus $(a,g)\in E_S$ which proves that $\operatorname{Ker}\varphi\subseteqq E_S$. This completes the proof of (i), (ii), (vi) and (vii).

(iii) We have $(a,g)\mathscr{R}(b,h)$ if and only if there exist inverses $(a,g)'$ and $(b,h)'$ of (a,g) and (b,h), respectively, such that $(a,g)(a,g)'\mathscr{R}(b,h)(b,h)'$ in E_S. By applying (ii) we deduce that this holds if and only if there exist $a^+\in V(g^{-1}a)$ and $b^+\in V(h^{-1}b)$ such that $a\circ ga^+\mathscr{R}b\circ hb^+$ in $E_{\mathscr{Y}}$. Since $a\mathscr{R}a\circ ga^+$ and $b\mathscr{R}b\circ hb^+$ in $\mathscr{X}$, this is equivalent to requiring that $a\mathscr{R}b$.

(iv) is proved dually to (iii).

(v) is an immediate consequence of (ii). The proof is complete.

In the terminology introduced here the results of Section 2 can be formulated in such a way that the triple $(G_S, \mathscr{X}, \mathscr{Y})$ defined there is a POM-triple and S is isomorphic to $PO(G_S, \mathscr{X}, \mathscr{Y})$. Thus we deduce the following

Proposition 3.3. *Every E-unitary regular semigroup is isomorphic to a PO-semigroup defined by a POM-triple.*

It is clear that, for a given PO-triple $(G, \mathscr{X}, \mathscr{Y})$, the PO-semigroup $PO(G, \mathscr{X}, \mathscr{Y})$ is a regular subsemigroup without 0 in the 0-semidirect product $\check{\mathscr{X}}*_0G$.

In the sequel we investigate the connection between PO-semigroups and regular subsemigroups without 0 in 0-semidirect products of a strictly combinatorial semigroup by a group.

Lemma 3.4. *Let G be a group acting on a semigroup T with 0. Then*

(i) *the 0-semidirect product $T*_0G$ is a semigroup in which $E_{T*_0G}=\{(e,1)\colon e\in E_T\setminus 0\}\cup 0$ and $V_{T*_0G}((t,g))=\{(g^{-1}t', g^{-1})\colon t'\in V_T(t)\}$ for every $(t,g)\in T*_0G$;*

(ii) *$T*_0G$ is regular if and only if T is regular;*

(iii) *$T*_0G$ is orthodox if and only if T is orthodox, and in this case, E_{T*_0G} is isomorphic to E_T;*

(iv) *$T*_0G$ is categorical at 0 if and only if T is categorical at 0;*

(v) *if T is regular and categorical at 0 then $T*_0G$ is $E\setminus 0$-unitary if and only if T is $E\setminus 0$-unitary.*

Proof. Statements (i)—(iv) can be easily proved therefore they are left to the reader. In order to prove (v) it suffices to check by Proposition 1.4 that $E_{T*_0G}\setminus 0$ is a left unitary subset in $T*_0G$ if and only if $E_T\setminus 0$ is a left unitary subset in T. Suppose first that $E_{T*_0G}\setminus 0$ is a left unitary subset in $T*_0G$, and let $e\in E_T$, $t\in T$ be such that $et\in E_T\setminus 0$. Then we have $(e, 1)(t, 1)=(et, 1)\in E_{T*_0G}\setminus 0$ and $(e, 1)\in E_{T*_0G}\setminus 0$ by (i), which imply that $(t, 1)\in E_{T*_0G}\setminus 0$. Thus, again applying (i), we infer that $t\in E_T\setminus 0$. Conversely, suppose that $E_T\setminus 0$ is a left unitary subset in T and $(e, 1)\in E_{T*_0G}$, $(t, g)\in T*_0G$ with $(e, 1)(t, g)\in E_{T*_0G}\setminus 0$. Then, by (i), we obtain that $g=1$ and $e, et\in E_T\setminus 0$. Hence it follows that $t\in E_T\setminus 0$, that is, $(t, g)\in E_{T*_0G}\setminus 0$. The proof is complete.

Lemma 3.4 implies that if T is strictly combinatorial then $T*_0G$ is an $E\setminus 0$-unitary regular semigroup which is categorical at 0. Consequently, every regular subsemigroup without 0 in $T*_0G$ is E-unitary. Hence we obtain

Proposition 3.5. *Every regular subsemigroup without 0 in a 0-semidirect product of a strictly combinatorial semigroup by a group is E-unitary.*

Now we turn to investigating the connection between PO-semigroups defined by POM-triples and maximal subsemigroups without 0 in 0-semidirect products of strictly combinatorial semigroups by groups. First of all, we determine the maximal subsemigroups without 0 in a 0-semidirect product of a strictly combinatorial semigroup by a group.

Lemma 3.6. *Let G be a group acting on a strictly combinatorial semigroup T. Then, in $T*_0G$, the maximal subsemigroups without 0 are*

$$M_i=\{(t, g)\in T*_0G\colon\ t\varrho\mathcal{R}i\mathcal{L}(g^{-1}t)\varrho\} \tag{4}$$

where $i\in E_{T/\varrho}\setminus 0$, and every subsemigroup without 0 in $T*_0G$ is contained in M_i for a unique $i\in E_{T/\varrho}\setminus 0$.

Proof. Since, in an $\mathcal{H}$-trivial Brandt semigroup, the only subsemigroups without 0 are the singletons containing idempotents, it suffices to find a 0-restricted homomorphism ψ of $T*_0G$ onto an $\mathcal{H}$-trivial Brandt semigroup such that the inverse images of the idempotents are just the M_i's. We shall use for this purpose an $\mathcal{H}$-trivial Brandt semigroup $B(I)$ which is the image of T under some 0-restricted homomorphism $\varphi\colon T\to B(I)$ with $\ker\varphi=\varrho$. Since φ is 0-restricted, $t\in T\setminus 0$ implies $t\varphi\neq 0$. Denote by φ_n $(n=1, 2)$ the mapping of $T\setminus 0$ into I assigning the nth component of $t\varphi$ to t for each $t\in T\setminus 0$. Define the mapping $\psi\colon T*_0G\to B(I)$ by $0\psi=0$ and $(t, g)\psi=[t\varphi_1, (g^{-1}t)\varphi_2]$. We prove that ψ is a 0-restricted homomorphism. By definition, ψ is 0-restricted. Now let $(t, g), (u, h)\in T*_0G$. Observe that $(t, g)(u, h)=0$ if and only if $t\cdot gu=0$, that is, if and only if $g^{-1}t\cdot u=0$. Since

φ is 0-restricted, the latter equality is equivalent to $(g^{-1}t\cdot u)\varphi=0$. This holds if and only if $(g^{-1}t)\varphi_2\neq u\varphi_1$. Thus we see that if $(t,g)(u,h)=0$ then $(t,g)\psi\cdot\cdot(u,h)\psi=0$. Moreover, if $(t,g)(u,h)\neq 0$ then $(g^{-1}t)\varphi_2=u\varphi_1$ and hence

$$((t,g)(u,h))\psi=(t\cdot gu, gh)\psi=\left[(t\cdot gu)\varphi_1,((gh)^{-1}(t\cdot gu))\varphi_2\right]=$$
$$=[t\varphi_1,(h^{-1}u)\varphi_2]=[t\varphi_1,(g^{-1}t)\varphi_2]\cdot[u\varphi_1,(h^{-1}u)\varphi_2]=(t,g)\psi\cdot(u,h)\psi.$$

Here we have utilized that $(t\cdot gu)\varphi=t\varphi\cdot(gu)\varphi$ implies $(t\cdot gu)\varphi_1=t\varphi_1$ and, similarly, $((gh)^{-1}(t\cdot gu))\varphi=((gh)^{-1}t\cdot h^{-1}u)\varphi=((gh)^{-1}t)\varphi\cdot(h^{-1}u)\varphi$ imply

$$((gh)^{-1}(t\cdot gu))\varphi_2=(h^{-1}u)\varphi_2.$$

The proof is complete.

Proposition 3.7. (i) *Let G be a group acting on a strictly combinatorial semigroup T. Let $i\in E_{T/\varrho}\setminus 0$. Define $\bar G=\{g\in G\colon t\varrho\mathcal{R}i\mathcal{L}(g^{-1}t)\varrho$ for some $t\in T\setminus 0\}$, $\mathcal{Y}=\{t\in T\setminus 0\colon t\varrho\mathcal{R}i\mathcal{L}(g^{-1}t)\varrho$ for some $g\in G\}$ and $\mathcal{X}=\{ga\colon g\in\bar G,\ a\in\mathcal{Y}\}$. Define a partial operation on $\mathcal{X}$ by restricting the operation of T to $\mathcal{X}$ and define an action of $\bar G$ on $\mathcal{X}$ by restricting the action of G on T to $\bar G$ and $\mathcal{X}$. Then $(\bar G,\mathcal{X},\mathcal{Y})$ is a POM-triple and $PO(\bar G,\mathcal{X},\mathcal{Y})=M_i$ (cf. (4)).*

(ii) *Conversely, for every POM-triple $(G,\mathcal{X},\mathcal{Y})$, the PO-semigroup $PO(G,\mathcal{X},\mathcal{Y})$ is a maximal subsemigroup without 0 in $\check{\mathcal{X}}*_0G$.*

Proof. (i) First we show that $\bar G$ is a subgroup in G. If $g\in G$, $t\in T\setminus 0$ with $t\varrho\mathcal{R}i\mathcal{L}(g^{-1}t)\varrho$ and $t'\in V_T(t)$, then we have $g^{-1}t'\in V_T(g^{-1}t)$ and $t'\varrho\mathcal{L}i\mathcal{R}(g^{-1}t')\varrho$. The latter relation can be written in the form $(g^{-1}t')\varrho\mathcal{R}i\mathcal{L}(g(g^{-1}t'))\varrho$. Since T is regular, this shows that $g\in\bar G$ implies $g^{-1}\in\bar G$. Assume that $g,h\in\bar G$. Then, by definition, there exist $t,u\in T\setminus 0$ with $t\varrho\mathcal{R}i\mathcal{L}(g^{-1}t)\varrho$ and $u\varrho\mathcal{R}i\mathcal{L}(h^{-1}u)\varrho$. Thus, by (4), we have $(t,g),(u,h)\in M_i$. Making use of the fact that, by Lemma 3.6, M_i is a subsemigroup in $T*_0G$, we obtain that $(t,g)(u,h)\in M_i$. Hence we infer that $(t\cdot gu)\varrho\mathcal{R}i\mathcal{L}((gh)^{-1}(t\cdot gu))\varrho$ which implies that $gh\in\bar G$. Thus $\bar G$ is, indeed, a subgroup in G.

Now we verify that $\mathcal{Y}$ is a right ideal in $\mathcal{X}$. Let $a\in\mathcal{Y}$ and $x\in\mathcal{X}$ such that $ax\neq 0$ in T. Then we have $(ax)\varrho\mathcal{R}a\varrho\mathcal{R}i$. On the other hand, by the definitions of $\mathcal{X}$ and $\mathcal{Y}$, there exist $g\in\bar G$ and $b\in\mathcal{Y}$ with $x=gb$, and $i\mathcal{L}(h^{-1}b)\varrho$ for some $h\in G$. Clearly, $h\in\bar G$ and $((gh)^{-1}(ax))\varrho=((gh)^{-1}a\cdot h^{-1}b)\varrho\mathcal{L}(h^{-1}b)\varrho\mathcal{L}i$. Thus, indeed, ax belongs to $\mathcal{Y}$ provided $a\in\mathcal{Y}$, $x\in\mathcal{X}$ and $ax\neq 0$ in T. This implies that $\check{\mathcal{X}}$ is a subsemigroup in T. For, let $x,y\in\mathcal{X}$ such that $xy\neq 0$ in T. Suppose that $x=gb$ where $g\in\bar G$ and $b\in\mathcal{Y}$. Since $\mathcal{Y}$ is a right ideal in $\mathcal{X}$, we have $b\cdot g^{-1}y\in\mathcal{Y}$ whence $xy=g(b\cdot g^{-1}y)\in\mathcal{X}$.

Now we show that $\check{\mathcal{X}}$ is regular. Let $ga\in\mathcal{X}$ where $g\in\bar G$ and $a\in\mathcal{Y}$. Then $a\varrho\mathcal{R}i\mathcal{L}(h^{-1}a)\varrho$ for some $h\in\bar G$. If $a'\in V_T(a)$ then we have $(h^{-1}a')\varrho\mathcal{R}i\mathcal{L}a'\varrho$

which implies that $h^{-1}a' \in \mathcal{Y} \cap V_T(h^{-1}a)$. Hence we deduce that $ga' = (gh)(h^{-1}a')$ is an inverse of ga in $\mathcal{X}$ for every $a' \in V_T(a)$. Consequently,

$$(5) \qquad V_T(x) = V_{\mathcal{X}}(x) \quad \text{for any} \quad x \in \mathcal{X}.$$

In particular, we obtain that $\check{\mathcal{X}}$ is regular. Finally, we verify that $\check{\mathcal{X}}/\varrho_{\check{\mathcal{X}}}$ is an $\mathcal{H}$-trivial Brandt semigroup. It is easy to see by definition that $\mathcal{X} = \{t \in T\colon (ge)\varrho \mathcal{R} t\varrho \mathcal{L} (he)\varrho$ for some $g, h \in \bar{G}\}$ where $e \in E_T$ with $e\varrho = i$. Hence $\mathcal{X}$ is a union of ϱ-classes in T and $\check{\mathcal{X}}\varrho$, which is isomorphic to $\check{\mathcal{X}}/\varrho_{\check{\mathcal{X}}}$, is an $\mathcal{H}$-trivial Brandt semigroup since T/ϱ is an $\mathcal{H}$-trivial Brandt semigroup. Thus we have shown that $\check{\mathcal{X}}$ is a strictly combinatorial semigroup.

Returning to the properties of $\mathcal{Y}$ observe that $E_{\mathcal{Y}} = \{e \in E_T \setminus 0\colon e\varrho = i\}$. This clearly implies that $E_{\mathcal{Y}}$ is a subband in $\mathcal{Y}$. Property (PO3) follows by the definition of $\mathcal{X}$ and that of the action of $\bar{G}$ on $\mathcal{X}$ while (PO4) is a consequence of the definitions of $\bar{G}$ and $\mathcal{Y}$ and of the fact (easily deduced from (5)) that

$$(6) \qquad V_{\mathcal{X}}(\mathcal{Y}) = V_T(\mathcal{Y}) = \{t \in T \setminus 0\colon t\varrho \mathcal{L} i \mathcal{R} (g^{-1}t)\varrho \text{ for some } g \in \bar{G}\}.$$

Thus $(\bar{G}, \mathcal{X}, \mathcal{Y})$ is a PO-triple.

Now we show that (M) is satisfied. Suppose that $(\bar{G}, \mathcal{X}, \mathcal{Y}_1)$ is a PO-triple and $\mathcal{Y} \subseteq \mathcal{Y}_1$. Since $\mathcal{Y}_1$ is a right ideal and $\mathcal{X}$ is regular, $y \in \mathcal{Y}_1$ implies $yy' \in E_{\mathcal{Y}_1}$ for any $y' \in V_{\mathcal{X}}(y)$. Therefore we obtain that $y\varrho \mathcal{R} (yy')\varrho = i$ as $E_{\mathcal{Y}_1}$ is a subband in $\mathcal{X} \subseteq T \setminus 0$ containing $E_{\mathcal{Y}}$. Assume that $y = ga$ for some $g \in \bar{G}$ and $a \in \mathcal{Y}$. Then there exists $h \in \bar{G}$ with $(h^{-1}a)\varrho \mathcal{L} i$ whence we obtain that $((gh)^{-1}y)\varrho = (h^{-1}a)\varrho \mathcal{L} i$. Thus $y \in \mathcal{Y}$ is proved. Hence $\mathcal{Y} = \mathcal{Y}_1$ and therefore $(\bar{G}, \mathcal{X}, \mathcal{Y})$ is a POM-triple.

By the definitions of $(\bar{G}, \mathcal{X}, \mathcal{Y})$ and the PO-semigroup $PO(\bar{G}, \mathcal{X}, \mathcal{Y})$, Proposition 3.2 (i) implies that $PO(\bar{G}, \mathcal{X}, \mathcal{Y})$ is a subsemigroup without 0 in $T *_0 G$. Then it follows from Lemma 3.6 that $PO(\bar{G}, \mathcal{X}, \mathcal{Y}) \subseteq M_i$. The reverse inclusion follows if we observe that $t\varrho \mathcal{R} i \mathcal{L} (g^{-1}t)\varrho$ implies that $t \in \mathcal{Y}$, $g \in \bar{G}$ and, by (6), we have $g^{-1}t \in V_{\mathcal{X}}(\mathcal{Y})$. The proof of the direct part is complete.

Now we turn to the proof of the converse part.

(ii) By Proposition 3.2 (i), $S = PO(G, \mathcal{X}, \mathcal{Y})$ is clearly a subsemigroup in $\check{\mathcal{X}} *_0 G$ and $0 \notin S$. Then, by Lemma 3.6, we have $S \subseteq M_i$ for a unique $i \in E_{\check{\mathcal{X}}/\varrho} \setminus 0$ and thus, by Lemma 3.1, we infer $\mathcal{Y} \subseteq \{x \in \mathcal{X}\colon x\varrho \mathcal{R} i\}$. The latter subset which we will denote by $\mathcal{Y}_i$, is easily seen to be a right ideal in $\mathcal{X}$ where $E_{\mathcal{Y}_i}$ is a subband. Since $(G, \mathcal{X}, \mathcal{Y})$ is assumed to be a POM-triple we infer that $\mathcal{Y} = \mathcal{Y}_i$. Hence, if $(a, g) \in M_i$ then $a \in \mathcal{Y}$ and $(g^{-1}a)\varrho \mathcal{L} i$ in $\check{\mathcal{X}}/\varrho$. The latter relation implies $b\varrho \mathcal{R} i$ for any $b \in V(g^{-1}a)$, that is, we have $V(g^{-1}a) \subseteq \mathcal{Y}_i = \mathcal{Y}$. Hence it follows that $g^{-1}a \in V(\mathcal{Y})$ and we have $(a, g) \in S$. Thus the equality $S = M_i$ is proved.

We can summarize the results of Sections 2 and 3 as follows:

Theorem 3.8. *For a regular semigroup S the following conditions are equivalent to each other:*

(i) *S is E-unitary;*

(ii) *S is isomorphic to a PO-semigroup;*

(iii) *S is isomorphic to a PO-semigroup defined by a POM-triple;*

(iv) *S is a regular subsemigroup without* 0 *in a* 0*-semidirect product of a strictly combinatorial semigroup by a group;*

(v) *S is a maximal subsemigroup without* 0 *in a* 0*-semidirect product of a strictly combinatorial semigroup by a group.*

Proof. The implications (i)⇒(iii) and (ii)⇒(i) are stated in Propositions 3.3 and 3.2 (i), respectively. Since (iii)⇒(ii) is trivial, the conditions (i), (ii) and (iii) are equivalent to each other. Moreover, the implications (iii)⇒(v) and (iv)⇒(i) follow from Propositions 3.7 (ii) and 3.5, respectively. The proof is complete as (v)⇒(iv) is easily deduced from Propositions 3.7 (i) and 3.2 (i).

4. E-unitary regular semigroups and connected idempotent and regular categories

A theory generalizing McAlister's P-theorem for not necessarily regular semigroups with commuting idempotents was developed by Margolis and Pin in [4]. Although their terminology and methods are entirely different from ours, for inverse semigroups their main theorem says almost the same as our Theorem 3.8 (i), (v). After making a dictionary between the two terminologies we deduce the generalization of Theorem 4.1 [4] for E-unitary regular semigroups.

In this section the reader is assumed to be familiar with the paper [4]. The notions and notations of [4] are used without any reference.

Let C be a category. Then Mor (C) is a partial groupoid. Denote by $\check{\mathrm{Mor}}\,(C)$ the groupoid obtained from Mor (C) by adjoining a new symbol 0 and extending the operation as in Section 1. Clearly, $\check{\mathrm{Mor}}\,(C)$ is a semigroup which is categorical at 0.

The following proposition states that categories can be considered as certain semigroups with 0 together with a 0-restricted homomorphism into an $\mathscr{H}$-trivial Brandt semigroup.

Proposition 4.1. (i) *For every category C, the groupoid* $\check{\mathrm{Mor}}\,(C)$ *is a semigroup and the mapping φ_C: $\check{\mathrm{Mor}}\,(C)\to B(\mathrm{Ob}\,(C))$ defined by $0\varphi_C=0$ and $p\varphi_C=[u,v]$ provided $p\in\mathrm{Mor}\,(u,v)$ is a* 0*-restricted homomorphism onto a full subsemigroup of $B(\mathrm{Ob}\,(C))$ such that, for every $e,f\in E_{B(\mathrm{Ob}\,(C))}$, $e\varphi_C^{-1}$ is a monoid with identity 1_e and $1_e p=p1_f=p$ for any $p\in\mathrm{Mor}\,(C)$ with $e\mathscr{R}p\varphi_C\mathscr{L}f$.*

(ii) *Let* S *be a semigroup with* 0 *and* $\varphi\colon S\to B(I)$ *a* 0*-restricted homomorphism onto a full subsemigroup of* $B(I)$ *such that, for every* $e,f\in E_{B(I)}$, $e\varphi^{-1}$ *is a monoid with identity* 1_e *and* $1_e s=s1_f=s$ *for any* $s\in S$ *with* $e\mathcal{R}s\varphi\mathcal{L}f$. *Then we can define a category* $C_{(S,\varphi)}$ *as follows:* $\mathrm{Ob}(C_{(S,\varphi)})=I$ *and* $\mathrm{Mor}\,(i,j)=\{s\in S\colon s\varphi=[i,j]\}$ *for any* $i,j\in I$.

(iii) *For every category* C, *we have* $C=C_{(\check{\mathrm{Mor}}(C),\varphi_C)}$.

(iv) *For every pair* (S,φ) *with properties required in* (ii), *we have* $S=\check{\mathrm{Mor}}\,C_{(S,\varphi)}$ *and* $\varphi=\varphi_{C_{(S,\varphi)}}$.

The proof is easy therefore it is left to the reader.

For brevity, a pair (S,φ) satisfying the properties required in (ii) is termed a *category pair.*

By the preceding proposition we immediately obtain

Proposition 4.2. (i) *The category* C *is connected if and only if* φ_C *is onto.*
(ii) *The category* C *is regular if and only if* $\check{\mathrm{Mor}}\,(C)$ *is regular.*
(iii) *The category* C *is idempotent if and only if* $\mathrm{Ker}\,\varphi_C\subseteq E_{\check{\mathrm{Mor}}(C)}$.

Assume that C is a connected, idempotent and regular category. Then the preceding proposition implies that $\check{\mathrm{Mor}}\,(C)$ is a regular semigroup and φ_C is a 0-restricted homomorphism of $\check{\mathrm{Mor}}\,(C)$ onto an $\mathcal{H}$-trivial Brandt semigroup with $\mathrm{Ker}\,\varphi_C\subseteq E_{\check{\mathrm{Mor}}(C)}$. Then Lemma 1.6 ensures that $\check{\mathrm{Mor}}\,(C)$ is $E\setminus 0$-unitary and $\ker\varphi_C=\varrho$. Thus $\check{\mathrm{Mor}}\,(C)$ is a strictly combinatorial semigroup in which every non-zero idempotent ϱ-class e is a monoid with identity 1_e and, for arbitrary non-zero idempotent ϱ-classes e,f and for any $p\in\mathrm{Mor}\,(C)$ with $e\mathcal{R}p\varrho\mathcal{L}f$, we have $1_e p=p1_f=p$. Now let S be a strictly combinatorial semigroup in which every non-zero idempotent ϱ-class e is a monoid with identity 1_e and, for any non-zero idempotent ϱ-classes e,f and for any $s\in S$ with $e\mathcal{R}s\varrho\mathcal{L}f$, we have $1_e s=s1_f=s$. Such an S will be termed a *strictly combinatorial semigroup with local identities.*

Let S be a strictly combinatorial semigroup with local identities. By definition, there exists an $\mathcal{H}$-trivial Brandt semigroup $B(I)$ and a surjective 0-restricted homomorphism $\varphi\colon S\to B(I)$ with $\ker\varphi=\varrho$. Clearly, (S,φ) is a category pair and, by Propositions 4.1 and 4.2, $C_{(S,\varphi)}$ is a connected, idempotent and regular category. Since Lemma 1.7 implies ϱ to be the only 0-restricted primitive inverse semigroup congruence on S, for each category pair (S,φ'), we have $\ker\varphi'=\varrho$. Consequently, for any category pairs $(S,\varphi\colon S\to B(I))$ and $(S,\varphi'\colon S\to B(I'))$, there exists an isomorphism $\psi\colon B(I)\to B(I')$ with $\varphi\psi=\varphi'$. Hence we can easily deduce

Proposition 4.3. *If* C *is a connected, idempotent and regular category then* $\check{\mathrm{Mor}}\,(C)$ *is a strictly combinatorial semigroup with local identities. Conversely, for every strictly combinatorial semigroup* S *with local identities, there exists an, up to* *isomorphisms, unique category* C_S *such that* $\check{\mathrm{M}}\mathrm{or}\,(C_S)$ *is isomorphic to* S. *The category* C_S *is connected, idempotent and regular.*

Recall that an *isomorphism of the category* C *onto the category* D is a functor $F\colon C\to D$ which induces a bijection of $\mathrm{Ob}\,(C)$ onto $\mathrm{Ob}\,(D)$ and a bijection of $\mathrm{Mor}\,(u, v)$ onto $\mathrm{Mor}\,(Fu, Fv)$ for every $u, v\in\mathrm{Ob}\,(C)$.

The connection between automorphisms of categories and automorphisms of the corresponding semigroups is easily described. Given a category C or a category pair (S, φ), denote by $\mathrm{Aut}\,C$ and $\mathrm{Aut}_\varphi S$, respectively, the group of automorphisms of C and the group of those automorphisms of S which possess the property that, for every $s, t\in S$, we have $s \ker\varphi\, t$ if and only if $s\alpha \ker\varphi\, t\alpha$.

Proposition 4.4. (i) *Let* $F\colon C\to C$ *be an automorphism of the category* C. *Then the mapping* F_m *induced by* F *on* $\mathrm{Mor}\,(C)$ *is an automorphism of the partial semigroup* $\mathrm{Mor}\,(C)$ *which can be extended to an automorphism of* $\check{\mathrm{M}}\mathrm{or}\,(C)$ *by setting* $0F_m=0$.

(ii) *Let* $(S, \varphi\colon S\to B(I))$ *be a category pair and* $\alpha\in\mathrm{Aut}_\varphi S$. *Then* α *induces an automorphism of* $B(I)$ *and, consequently, a permutation* π_α *of* I *in such a way that, if* $s\varphi=[i,j]$ *for some* $s\in S\setminus 0$ *and* $i,j\in I$, *then* $(s\alpha)\varphi=[i\pi_\alpha, j\pi_\alpha]$. *Define a functor* $F_\alpha\colon C_{(S,\varphi)}\to C_{(S,\varphi)}$ *as follows:* $F_\alpha i=i\pi_\alpha$ *for every* $i\in I=\mathrm{Ob}\,(C_{(S,\varphi)})$ *and* $F_\alpha s=s\alpha$ *for every* $s\in\hat{S}=\mathrm{Mor}\,(C_{(S,\varphi)})$. *Then* $F_\alpha\in\mathrm{Aut}\,C_{(S,\varphi)}$.

(iii) *The mappings* $(\mathrm{Aut}\,C)^d\to\mathrm{Aut}_{\varphi_C}\check{\mathrm{M}}\mathrm{or}\,(C)$, $F\mapsto F_m$ *and* $\mathrm{Aut}_{\varphi_C}\check{\mathrm{M}}\mathrm{or}\,(C)\to(\mathrm{Aut}\,C)^d$, $\alpha\mapsto F_\alpha$ *defined in* (i) *and* (ii) *are group-isomorphisms inverse to each other.*

(iv) *The mappings* $\mathrm{Aut}_\varphi S\to(\mathrm{Aut}\,C_{(S,\varphi)})^d$, $\alpha\mapsto F_\alpha$ *and* $(\mathrm{Aut}\,C_{(S,\varphi)})^d\to\mathrm{Aut}_\varphi S$, $F\mapsto F_m$ *are group-isomorphisms inverse to each other.*

By applying this description for connected, idempotent and regular categories the case becomes simpler. For, if S is a strictly combinatorial semigroup and (S, φ) is a category pair, then we have seen that $\ker\varphi=\varrho$. By Proposition 1.5, it is easy to check that, for any $s, t\in S\setminus 0$, we have $s\varrho t$ if and only if $s\alpha\varrho t\alpha$. Hence $\mathrm{Aut}_\varphi S=\mathrm{Aut}\,S$, the group of all automorphisms of S.

Proposition 4.5. (i) *If* C *is a connected, idempotent and regular category, then the mapping* $(\mathrm{Aut}\,C)^d\to\mathrm{Aut}\,\check{\mathrm{M}}\mathrm{or}\,(C)$, $F\mapsto F_m$ (cf. Proposition 4.4) *is a group-isomorphism.*

(ii) *If* S *is a strictly combinatorial semigroup with local identities and* C_S *is a connected, idempotent and regular category with* $\mathrm{Mor}\,(C_S)=\hat{S}$, *then, for every* $\alpha\in\mathrm{Aut}\,S$, *there exists a unique functor* F_α *which coincides with* α *on* $\mathrm{Mor}\,(C_S)$. *Moreover, the mapping* $\mathrm{Aut}\,S\to(\mathrm{Aut}\,C_S)^d$, $\alpha\mapsto F_\alpha$ (cf. Proposition 4.4) *is a group-isomorphism.*

Remark. Proposition 4.5 implies that if a group G acts on a connected, idempotent and regular category C then this action determines in a natural way an action of G on $\check{\mathrm{Mor}}(C)$. Conversely, if a group G acts on a strictly combinatorial semigroup S with local identities and C_S is a category with $\mathrm{Mor}(C_S)=\hat{S}$, then this action can be extended to an action of G on C_S.

It remains to find connection between the monoids C_u (defined in [4]) and PO-semigroups.

Proposition 4.6. (i) *Let G be a group and C a connected, idempotent and regular category on which G acts transitively, and let $u\in\mathrm{Ob}(C)$. Then $(G, \mathrm{Mor}(C), \mathcal{Y}_u)$ with $\mathcal{Y}_u=\mathrm{Mor}(u, C)$ is a POM-triple and $C_u=PO(G, \mathrm{Mor}(C), \mathcal{Y}_u)$.*

(ii) *Let $(G, \mathcal{X}, \mathcal{Y})$ be a POM-triple where $\check{\mathcal{X}}$ is a strictly combinatorial semigroup with local identities. Let $C_{\mathcal{X}}$ be a category such that $\mathrm{Mor}(C_{\mathcal{X}})=\mathcal{X}$. Then $C_{\mathcal{X}}$ is a connected, idempotent and regular category on which G acts transitively, and $PO(G, \mathcal{X}, \mathcal{Y})=(C_{\mathcal{X}})_u$ for some $u\in\mathrm{Ob}(C_{\mathcal{X}})$.*

Proof. (i) By Proposition 4.3 and Remark, $\check{\mathrm{Mor}}(C)$ is a strictly combinatorial partial semigroup on which G acts. Let us define $\bar{G}$, $\mathcal{Y}$ and $\mathcal{X}$ by means of G, $T=\check{\mathrm{Mor}}(C)$ and $[u, u]\in E_{(\check{\mathrm{Mor}}(C))\varphi_C}$ as in Proposition 3.7 (i). We claim that $\bar{G}=G$, $\mathcal{Y}=\mathcal{Y}_u$ and $\mathcal{X}=\mathrm{Mor}(C)$.

First of all, observe that, for any $g\in G$ and $p\in\mathrm{Mor}(C)$, we have

$$p\in\mathrm{Mor}(u, gu) \quad \text{if and only if} \quad p\varphi_C\mathcal{R}[u, u]\mathcal{L}(g^{-1}p)\varphi_C. \tag{7}$$

On the one hand, hence it follows that $\bar{G}=G$ as C is connected and therefore $\mathrm{Mor}(u, gu)\neq\square$ for every $g\in G$. On the other hand, we obtain from (7) the equality $\mathcal{Y}=\mathcal{Y}_u$ by making use of the assumption that G acts transitively, and therefore any $v\in\mathrm{Ob}(C)$ is of the form gu for some $g\in G$. Now let $q\in\mathrm{Mor}(v, C)$. As we have just seen, there exists $g\in G$ with $gu=v$. Since g induces an automorphism on C, we infer that there exists $p\in\mathrm{Mor}(u, C)$ with $gp=q$. Hence $\mathcal{X}=G\mathcal{Y}_u=\mathrm{Mor}(C)$. Thus Proposition 3.7 (i) ensures that $(G, \mathrm{Mor}(C), \mathcal{Y}_u)$ is a POM-triple. The equality $C_u=PO(G, \mathrm{Mor}(C), \mathcal{Y}_u)$ immediately follows as $V(\mathcal{Y}_u)=\mathrm{Mor}(C, u)$.

(ii) Proposition 4.3 implies $C_{\mathcal{X}}$ to be a connected, idempotent and regular category. In the proof of Proposition 3.7 (ii) it is verified that $\mathcal{Y}=\{x\in\mathcal{X}: x\varrho\mathcal{R}i\}$ for some $i\in E_{\check{\mathcal{X}}/\varrho}$. Hence it follows that $\mathcal{Y}=\mathrm{Mor}(u, C_{\mathcal{X}})$ for some $u\in\mathrm{Ob}(C_{\mathcal{X}})$. Now let $v, w\in\mathrm{Ob}(C_{\mathcal{X}})$. Since $C_{\mathcal{X}}$ is connected, there exist $x\in\mathrm{Mor}(v, C_{\mathcal{X}})$ and $y\in\mathrm{Mor}(w, C_{\mathcal{X}})$. Since $G\mathcal{Y}=\mathcal{X}$ we have $g, h\in G$ and $a, b\in\mathcal{Y}$ with $ga=x$ and $hb=y$. Then the action of G on $C_{\mathcal{X}}$ has the property that $gu=v$ and $hu=w$. Hence we infer that $hg^{-1}v=w$, that is, G acts transitively on $C_{\mathcal{X}}$. Thus $C_{\mathcal{X}}$ satisfies the conditions required in (i) whence it follows that $(C_{\mathcal{X}})_u=PO(G, \mathrm{Mor}(C_{\mathcal{X}}), \mathcal{Y}_u)$ where $\mathrm{Mor}(C_{\mathcal{X}})=\mathcal{X}$ and $\mathcal{Y}_u=\mathcal{Y}$. This completes the proof.

Now we are ready to give a condition equivalent to each of (i)—(v) in Theorem 3.8 which is analogous to that in the main theorem of [4].

Theorem 4.7. *The following condition is equivalent to each of the conditions* (i)—(v) *in Theorem* 3.8.

(vi) S^1 *is isomorphic to a monoid* C/G *where* G *is a group acting transitively without fixpoints on a connected, idempotent and regular category* C.

Proof. Let S be an E-unitary regular semigroup. Then S^1 is also an E-unitary regular semigroup. By the method described in Section 2 we can construct a POM-triple $(G_{S^1}, \mathscr{X}, \mathscr{Y})$ (cf. Proposition 3.3). Consider the pair $(\check{\mathscr{X}}, \varphi)$ where φ is the homomorphism defined in (II) of Section 2. By (I) and (II), $\check{\mathscr{X}}$ is a strictly combinatorial semigroup and φ is a 0-restricted homomorphism of $\check{\mathscr{X}}$ onto $B(G_{S^1})$. It is easy to check by (1) that $(1, g)\circ(s, g)=(s, g)\circ(1, g\cdot s\sigma)=(s, g)$ for every $s\in S^1$ and $g\in G_{S^1}$. Since $(1, g)\varphi=[g, g]$, this implies that $\check{\mathscr{X}}$ is a strictly combinatorial semigroup with local identities. Thus $(\check{\mathscr{X}}, \varphi)$ is a category pair. Then, by Proposition 4.1 (ii), $C_{(\check{\mathscr{X}},\varphi)}$ is a category with $\mathrm{Mor}\,(C_{(\check{\mathscr{X}},\varphi)})=\mathscr{X}$. Moreover, Proposition 4.6 (ii) ensures that $C_{(\check{\mathscr{X}},\varphi)}$ is connected, idempotent and regular, and G_{S^1} acts on it transitively. Observe that the automorphism of $C_{(\check{\mathscr{X}},\varphi)}$ determined by an element $g\in G_{S^1}$ induces the regular left translation on $G_{S^1}=\mathrm{Ob}\,(C_{(\check{\mathscr{X}},\varphi)})$ corresponding to $g\in G_{S^1}$. Thus G_{S^1} acts on $C_{(\check{\mathscr{X}},\varphi)}$ without fixpoints. Property (VI) in Section 2 ensures S^1 to be isomorphic to $PO(G_{S^1}, \mathscr{X}, \mathscr{Y})$, and Proposition 4.6 (ii) implies that $PO(G_{S^1}, \mathscr{X}, \mathscr{Y})=(C_{(\check{\mathscr{X}},\varphi)})_u$ for some $u\in\mathrm{Ob}\,(C_{(\check{\mathscr{X}},\varphi)})$. Hence S^1 is isomorphic to $(C_{(\check{\mathscr{X}},\varphi)})_u$ for some $u\in\mathrm{Ob}\,(C_{(\check{\mathscr{X}},\varphi)})$. To complete the proof of the implication (i)$\Rightarrow$(vi) we refer to Proposition 3.11 [4] which states that if C is a category on which a group G acts transitively without fixpoints, then, for all $u\in\mathrm{Ob}\,(C)$, the monoid C_u is isomorphic to C/G.

Conversely, suppose G is a group acting transitively without fixpoints on a connected, idempotent and regular category C. Then Proposition 3.11 [4] just cited implies that C/G is isomorphic to C_u for all $u\in\mathrm{Ob}\,(C)$, while Proposition 4.6 (i) ensures C_u to be a PO-semigroup. Thus (vi) implies (iii), completing the proof of the theorem.

Finally, we show how one can reobtain McAlister's P-theorem from our results. Let $(G, \mathscr{X}, \mathscr{Y})$ be a PO-triple such that $PO(G, \mathscr{X}, \mathscr{Y})$ is an inverse semigroup or, equivalently, $E_{\mathscr{Y}}$ is a semilattice. It is not difficult to check that $(G, \widehat{\check{\mathscr{X}}/\gamma}, \mathscr{Y}\gamma)$ is also a PO-triple. We claim that the mapping $\eta\colon PO(G, \mathscr{X}, \mathscr{Y})\to PO(G, \widehat{\check{\mathscr{X}}/\gamma}, \mathscr{Y}\gamma)$, $(a, g)\eta=(a\gamma, g)$ is an isomorphism. It is immediate that η is a homomorphism. Let us verify that η is one-to-one. Assume that $(a, g), (b, h)\in PO(G, \mathscr{X}, \mathscr{Y})$ with $(a, g)\eta=(b, h)\eta$. Then $g=h$ and $a\gamma b$ in $\check{\mathscr{X}}$. The latter relation implies that $V(a)=$ $=V(b)$ and hence $V(g^{-1}a)=V(g^{-1}b)$. Since $(a, g), (b, g)\in PO(G, \mathscr{X}, \mathscr{Y})$ we have $g^{-1}a, g^{-1}b\in V(\mathscr{Y})$. Therefore there exists $c\in V(g^{-1}a)=V(g^{-1}b)$ with $c\in\mathscr{Y}$. Thus $(c, g^{-1})\in PO(G, \mathscr{X}, \mathscr{Y})$ and (c, g^{-1}) is an inverse of both (a, g) and (b, g). Since, by assumption, $PO(G, \mathscr{X}, \mathscr{Y})$ is an inverse semigroup we obtain that $(a, g)=(b, h)$. Now we show that η is onto. Consider an arbitrary element $(x\gamma, g)$ in $PO(G, \widehat{\mathscr{X}/\gamma}, \mathscr{Y}\gamma)$ where $x\in\mathscr{X}$. Then $x\gamma\in\mathscr{Y}\gamma$ and $g^{-1}(x\gamma)\in V(\mathscr{Y}\gamma)$. The first relation implies the existence of an element a in $\mathscr{Y}$ with $a\gamma x$, and hence $g^{-1}x\gamma g^{-1}a$. Thus, by the second relation we see that $g^{-1}a\in V(\mathscr{Y})$ since $V(\mathscr{Y})=V(\{x\in\mathscr{X}\colon x\gamma b \text{ for some } b\in\mathscr{Y}\})$. Therefore $(x\gamma, g)=(a, g)\eta$ which completes the proof of the fact that η is an isomorphism.

The strictly combinatorial semigroup $\check{\mathscr{X}}/\gamma$ is an inverse semigroup. Thus, by slightly modifying the proof of Theorem 4.2 [4], we can deduce the following assertion. The triple $(G, \bar{\mathscr{X}}, \bar{\mathscr{Y}})$ where $\bar{\mathscr{X}}$ is the partially ordered set of $\mathscr{J}$-classes of $\widehat{\mathscr{X}/\gamma}$, $\bar{\mathscr{Y}}=\{J\in\bar{\mathscr{X}}\colon J\cap E_{\mathscr{Y}\gamma}\neq\square\}$ and the action of G on $\bar{\mathscr{X}}$ is defined by $g(x\mathscr{J})=(gx)\mathscr{J}$ $(g\in G,\ x\in\widehat{\mathscr{X}/\gamma})$ is a McAlister triple and $PO(G, \widehat{\mathscr{X}/\gamma}, \mathscr{Y}\gamma)$ is isomorphic to $P(G, \bar{\mathscr{X}}, \bar{\mathscr{Y}})$. By the preceding paragraph this implies that $PO(G, \mathscr{X}, \mathscr{Y})$ is isomorphic to $P(G, \bar{\mathscr{X}}, \bar{\mathscr{Y}})$.

Consider the partially ordered set of $\mathscr{J}$-classes of $\mathscr{X}$ and denote it by $\tilde{\mathscr{X}}$. Put $\tilde{\mathscr{Y}}=\{J\in\tilde{\mathscr{X}}\colon J\cap E_{\mathscr{Y}}\neq\square\}$ and define an action of G on $\tilde{\mathscr{X}}$ by $g(x\mathscr{J})=(gx)\mathscr{J}$ for every $g\in G$ and $x\in\mathscr{X}$. Since $\gamma\subseteq\mathscr{J}$ on $\check{\mathscr{X}}$, it is easily seen that the mapping $\nu\colon\tilde{\mathscr{X}}\to\bar{\mathscr{X}}$ defined by $(x\mathscr{J})\nu=(x\gamma)\mathscr{J}$ $(x\in\mathscr{X})$ is an order isomorphism with the properties that $\tilde{\mathscr{Y}}\nu=\bar{\mathscr{Y}}$ and $g(\tilde{x}\nu)=(g\tilde{x})\nu$ for every $g\in G$ and $\tilde{x}\in\tilde{\mathscr{X}}$. Thus the triple $(G, \tilde{\mathscr{X}}, \tilde{\mathscr{Y}})$ is equivalent to the triple $(G, \bar{\mathscr{X}}, \bar{\mathscr{Y}})$. Therefore $(G, \tilde{\mathscr{X}}, \tilde{\mathscr{Y}})$ is also a McAlister triple and $P(G, \tilde{\mathscr{X}}, \tilde{\mathscr{Y}})$ is isomorphic to $P(G, \bar{\mathscr{X}}, \bar{\mathscr{Y}})$. The following theorem sums up what we have just proved.

Theorem 4.8. *Let $(G, \mathscr{X}, \mathscr{Y})$ be a PO-triple such that $E_{\mathscr{Y}}$ is a semilattice. Let $\tilde{\mathscr{X}}$ be the partially ordered set of $\mathscr{J}$-classes of $\mathscr{X}$ and $\tilde{\mathscr{Y}}=\{J\in\tilde{\mathscr{X}}\colon J\cap E_{\mathscr{Y}}\neq\square\}$. Define an action of G on $\tilde{\mathscr{X}}$ by $g(x\mathscr{J})=(gx)\mathscr{J}$ for every $g\in G$ and $x\in\mathscr{X}$. Then $(G, \tilde{\mathscr{X}}, \tilde{\mathscr{Y}})$ is a McAlister triple and $PO(G, \mathscr{X}, \mathscr{Y})$ is isomorphic to $P(G, \tilde{\mathscr{X}}, \tilde{\mathscr{Y}})$.*

References

[1] A. H. Clifford and G. B. Preston, *The algebraic theory of semigroups. I—II.*, Amer. Math. Soc. (Providence, R. I., 1961., 1967).

[2] J. M. Howie, *An introduction to semigroup theory*, Academic Press (London—New York—San Francisco, 1976).

[3] J. M. Howie and G. Lallement, Certain fundamental congruences on a regular semigroup, *Proc. Glasgow Math. Assoc.*, 7 (1966), 145—159.

[4] S. W. Margolis and J. E. Pin, Inverse semigroups and extensions of groups by semilattices, *J. Algebra*, to appear.

[5] D. B. McAlister, Groups, semilattices and inverse semigroups, *Trans. Amer. Math. Soc.*, **192** (1974), 227—244.

[6] D. B. McAlister, Groups, semilattices and inverse semigroups. II, *Trans. Amer. Math. Soc.*, **196** (1974), 351—370.

[7] W. D. Munn, A note on E-unitary inverse semigroups, *Bull. London Math. Soc.*, **8** (1976), 71—76.

[8] L. O'Carroll, Embedding theorems for proper inverse semigroups, *J. Algebra*, **42** (1976), 26—40.

[9] T. Saitô, Ordered regular proper semigroups, *J. Algebra*, **8** (1968), 450—477.

[10] M. B. Szendrei, On a pull-back diagram for orthodox semigroups, *Semigroup Forum*, **20** (1980), 1—10; Corr. **25** (1982), 311—324.

[11] M. B. Szendrei, On Takizawa's construction for E-unitary $\mathscr{R}$-unipotent semigroups, *Beitr. Algebra Geom.*, **20** (1985), 197—201.

[12] M. B. Szendrei, E-unitary $\mathscr{R}$-unipotent semigroups, *Semigroup Forum*, **32** (1985), 87—96.

[13] M. B. Szendrei, E-unitary regular semigroups, *Proc. Roy. Soc. Edinburgh Sect. A*, **106** (1987), 89—102.

[14] K. Takizawa, E-unitary $\mathscr{R}$-unipotent semigroups, *Bull. Tokyo Gakugei Univ.*, Ser. IV, **30** (1978), 21—33.

[15] K. Takizawa, Orthodox semigroups and E-unitary regular semigroups, *Bull. Tokyo Gakugei Univ.*, Ser. IV, **31** (1979), 41—43.

JÓZSEF ATTILA UNIVERSITY
BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Acta Sci. Math., 51 (1987), 251—268

Idempotent algebras with restrictions on subalgebras

ÁGNES SZENDREI

Dedicated to the memory of András P. Huhn

We study some consequences of an interesting property of idempotent algebras, namely, that their direct squares have enough reduced subalgebras in the following sense: For arbitrary idempotent algebra $\mathfrak{A}$, every reduced subalgebra $\mathfrak{B}$ of any finite power $\mathfrak{A}^n$ $(n>1)$ of $\mathfrak{A}$ produces reduced subalgebras in $\mathfrak{A}^2$, unless $\mathfrak{B}$ is a subdirect product of pairwise isomorphic, simple, locally affine subalgebras of $\mathfrak{A}$ (see Theorem 1.1). Section 1 contains also some applications. It follows that an idempotent algebra is locally quasi-primal if and only if it has no nonsingleton, locally affine subalgebras, and its square has no reduced subalgebras (Corollary 1.2). More generally, an idempotent algebra is locally para-primal if and only if its square has no reduced subalgebras (Corollary 1.3). For comparison, recall Rosenberg's Theorem [9] implying that in order to verify a finite algebra $\mathfrak{A}=(A; F)$ to be primal, one has to exclude the existence of certain types of subalgebras in $\mathfrak{A}^n$ with n running up to $n=|A|$.

In Section 2 we determine, up to local term equivalence, all idempotent algebras (of cardinality greater than 2) having no nonsingleton proper subalgebras (Theorem 2.1). They turn out to fall into three types: (a) locally quasi-primal algebras, (b) algebras locally term equivalent to the full idempotent reduct of a simple module, and (c) algebras whose clones of local term operations form a family of disjoint descending $(\omega+1)$-chains; these $(\omega+1)$-chains are related to "higher dimensional crosses" among the subalgebras of finite powers of the corresponding algebras.

This description is applied in Section 3 to finite algebras with minimal clones. It is proved that a finite algebra $(A; p)$ with p a Mal'tsev operation has a minimal clone if and only if p arises from an elementary Abelian group on A (Theorem 3.1); furthermore, a finite idempotent groupoid with minimal clone is term equivalent

Received June 30, 1986.

Research partially supported by Hungarian National Foundation for Scientific Research grant no. 1813.

to an algebra of this form if and only if it has a minimal nonsingleton subgroupoid of cardinality greater than 2 (Theorem 3.2).

The investigations in Sections 1 and 2 were partly inspired by the problem of determining the maximal subclones of the clone $\mathcal{I}_A$ of all idempotent operations on a finite set A, which was raised by I. G. Rosenberg during the Séminaire de Mathématiques Supérieures on "Universal algebra and relations" (Montreal, 1984), and was solved independently by several participants. Here the solution is derived from Theorem 1.1 (see Corollary 1.4 in the case when A is finite).

0. Preliminaries

For a nonempty set A, the clone of all operations on A will be denoted by $\mathcal{O}_A$, and for $n \geqq 1$, $\mathcal{O}_A^{(n)}$ will designate the set of n-ary operations on A. We write $|A|$ for the cardinality of A. Recall that an operation $f \in \mathcal{O}_A^{(n)}$ is said to *preserve* a subset B of A^k $(k \geqq 1)$ iff B is a subuniverse of the algebra $(A; f)^k$. For arbitrary mapping $g\colon A_1 \times \ldots \times A_n \to A_{n+1}$ $(n \geqq 1,\ A_1, \ldots, A_{n+1} \subseteq A)$ we define a subset $g_\square$ of A^{n+1} by

$$g_\square = \{(g(x_1, \ldots, x_n), x_1, \ldots, x_n)\colon x_1 \in A_1, \ldots, x_n \in A_n\}.$$

For arbitrary operations $f, f' \in \mathcal{O}_A$, f is said to *commute* with f' iff f preserves $(f')_\square$. It is easy to see that the commutativity of operations is a symmetric relation. If $\pi\colon B \to C$ is a bijection [or $\pi\colon A \to A$ is a permutation], then $\pi_\square$ will also be called a bijection [or permutation, respectively].

We now introduce some notation for constructions that will be used to produce subuniverses from subuniverses. Let B be a subset of A^k $(k \geqq 1)$. We will write $\mathbf{k}$ for the set $\{1, \ldots, k\}$ indexing the components of B. For an l-tuple $(i_1, \ldots, i_l) \in \mathbf{k}$ we define the projection of B onto its components $i_1, \ldots, i_l$ by

$$\mathrm{pr}_{i_1, \ldots, i_l} B = \{(x_{i_1}, \ldots, x_{i_l})\colon (x_1, \ldots, x_k) \in B\}.$$

In particular, if $l = k$ and $i_1, \ldots, i_k$ is a permutation of $1, \ldots, k$, then $\mathrm{pr}_{i_1, \ldots, i_k} B$ arises from B by rearranging the components. The property that, up to the order of their components, the subsets B and B' of A^k coincide, will be denoted by $B \approx B'$. For a nonvoid subset I of $\mathbf{k}$ with $I = \{i_1, \ldots, i_l\}$, $i_1 < \ldots < i_l$, we write $\mathrm{pr}_I B$ for $\mathrm{pr}_{i_1, \ldots, i_l} B$. The symbol $B \leqq B_1 \times \ldots \times B_k$ will be used to designate that $\mathrm{pr}_i B = B_i$ for all $i \in \mathbf{k}$. For $B \leqq B_1 \times \ldots \times B_k$ and for arbitrary bijections $\pi_i\colon B_i \to C_i$ $(C_i \subseteq A$, $i \in \mathbf{k})$ we set

$$B[\pi_1, \ldots, \pi_k] = \{(x_1\pi_1, \ldots, x_k\pi_k)\colon (x_1, \ldots, x_k) \in B\}.$$

If $1 \leqq l \leqq k$ and $(a_{l+1}, \ldots, a_k) \in A^{k-l}$, then we define the subset of A^l arising from B by "substituting the constants $a_{l+1}, \ldots, a_k$ for the $(l+1)$-st up to the k-th components" as follows:

$$B(x_1, \ldots, x_l, a_{l+1}, \ldots, a_k) = \{(x_1, \ldots, x_l)\in A^l\colon (x_1, \ldots, x_l, a_{l+1}, \ldots, a_k)\in B\}.$$

Let $\mathscr{C}$ be a clone on A. We say that an operation $g\in\mathcal{O}_A^{(n)}$ *can be interpolated* by operations from $\mathscr{C}$ iff for every finite subset S of A^n there exists an operation $f\in\mathscr{C}$ agreeing with g on S. The clone $\mathscr{C}$ is called *locally closed* iff it contains every operation that can be interpolated by its members. For an algebra $\mathfrak{A}=(A; F)$ the operations that can be interpolated by term [polynomial] operations of $\mathfrak{A}$ are called *local term [polynomial] operations* of $\mathfrak{A}$. It is easy to see that the local term [polynomial] operations of $\mathfrak{A}$ form a locally closed clone, which will be denoted by $\mathscr{T}_{\mathrm{loc}}(\mathfrak{A})$ [$\mathscr{P}_{\mathrm{loc}}(\mathfrak{A})$]; moreover, $\mathscr{T}_{\mathrm{loc}}(\mathfrak{A})$ [$\mathscr{P}_{\mathrm{loc}}(\mathfrak{A})$] is the least locally closed clone containing the clone $\mathscr{T}(\mathfrak{A})$ [resp., $\mathscr{P}(\mathfrak{A})$] of term [polynomial] operations of $\mathfrak{A}$. Clearly, if $\mathfrak{A}$ is finite, then $\mathscr{T}_{\mathrm{loc}}(\mathfrak{A})=\mathscr{T}(\mathfrak{A})$ and $\mathscr{P}_{\mathrm{loc}}(\mathfrak{A})=\mathscr{P}(\mathfrak{A})$. Two algebras with the same universe are called *term equivalent [locally term equivalent]* iff their clones of term operations [local term operations] coincide. It is well known that the algebras are determined, up to local term equivalence, by the subuniverses of their finite powers in the following sense: For an algebra $\mathfrak{A}=(A; F)$ and for $f\in\mathcal{O}_A$ we have $f\in\mathscr{T}_{\mathrm{loc}}(\mathfrak{A})$ if and only if f preserves every subuniverse of each finite power of $\mathfrak{A}$.

Let $\mathbf{A}=(A; +, -, 0)$ be an Abelian group. An algebra $\mathfrak{A}=(A; F)$ is said to be *affine [locally affine]* with respect to $\mathbf{A}$ iff the Mal'tsev operation $x-y+z$ is a term operation [resp., local term operation] of $\mathfrak{A}$, and every operation (hence every local term operation) of $\mathfrak{A}$ commutes with $x-y+z$. The Abelian group $\mathbf{A}$ is called *elementary* (or more precisely, an *elementary Abelian q-group*) iff for some prime q, $qa=0$ holds for all $a\in A$.

1. Reduced subalgebras of finite powers of idempotent algebras

Let A be a nonempty set. A subset B of A^k $(k\geq 1)$ is said to be *directly indecomposable* iff $B\not\approx(\mathrm{pr}_I B)\times(\mathrm{pr}_{\bar I} B)$ holds for all partitions $\{I, \bar I\}$ of $\mathbf{k}$, and B is *reduced* iff it is directly indecomposable and no projection $\mathrm{pr}_{i,j}B$ $(1\leq i<j\leq k)$ of B is a bijection. A subalgebra of some finite power of an algebra is called reduced iff its universe has this property. The *size* of B is the cardinal $\max\{|\mathrm{pr}_i B|\colon 1\leq i\leq k\}$.

The main result of this section is

1.1. Theorem. *Let $\mathfrak{A}=(A; F)$ be an idempotent algebra. For any $n\geq 2$ and for arbitrary reduced subuniverse $B\leq B_1\times\ldots\times B_n$ $(B_1, \ldots, B_n\subseteq A)$ of $\mathfrak{A}^n$ one of the following conditions holds:*

(1.1.1) *$\mathfrak{A}^2$ has a reduced subuniverse of the same size as B; or*

(1.1.2) *$\mathfrak{B}_i=(B_i; F)$ $(1\leq i\leq n)$ are isomorphic locally affine subalgebras of $\mathfrak{A}$, moreover, there exist a division ring K and a vector space ${}_K\mathbf{B}_1=(B_1; +, K)$ such* *that* $\mathfrak{B}_1$ *is locally term equivalent to the full idempotent reduct of the module* ${}_{(\mathrm{End}_K \mathbf{B}_1)}\mathbf{B}_1$. *For arbitrary isomorphisms* $\pi_i\colon \mathfrak{B}_i \to \mathfrak{B}_1$ $(1 \leqq i \leqq n)$, *the subuniverse* $B[\pi_1, \dots, \pi_n]$ *of* $\mathfrak{A}^n$ *has the form*

$$(1)\qquad B[\pi_1, \dots, \pi_n] \approx$$

$$\approx \{(y_1, \dots, y_{u-1}, g_u(y_1, \dots, y_{u-1}), \dots, g_n(y_1, \dots, y_{u-1}))\colon y_1, \dots, y_{u-1} \in B_1\}$$

for some $u \in \mathbf{n}$ $(u \geqq 2)$ *and for some operations* $g_u, \dots, g_n \in \mathscr{P}({}_K\mathbf{B}_1)$.

This is an extension of Theorem 4.3 (see also the remark following its proof) in [13] to not necessarily finite algebras. Before sketching the proof, which is quite similar to that of the finite version, we present several applications.

Theorem 1.1 yields nice criteria for idempotent algebras to be locally quasi-primal or para-primal, respectively. Recall that an algebra $\mathfrak{A} = (A; F)$ is called *locally quasi-primal,* or *quasi-primal* if $\mathfrak{A}$ is finite, iff every operation preserving all isomorphisms between subalgebras of $\mathfrak{A}$ is a local term operation of $\mathfrak{A}$ (A. F. Pixley [5], [6]). Equivalently, $\mathfrak{A}$ is locally quasi-primal iff $\mathfrak{A}^k$ has no reduced subuniverses for $k \geqq 2$ (see P. H. Krauss [2]). Combining the latter characterization with Theorem 1.1 we immediately get

1.2. Corollary. *An idempotent algebra* $\mathfrak{A} = (A; F)$ *is locally quasi-primal if and only if* $\mathfrak{A}$ *has no nonsingleton, locally affine subalgebras and* $\mathfrak{A}^2$ *has no reduced subalgebras.*

An algebra $\mathfrak{A} = (A; F)$ is called *locally para-primal* iff for every $k \geqq 1$, for every subuniverse B of $\mathfrak{A}^k$, and for every set $I \subseteq \mathbf{k}$ which is minimal with respect to the property that the projection $B \to \mathrm{pr}_I B$ is one-to-one, the equality $\mathrm{pr}_I B = \prod_{i \in I} \mathrm{pr}_i B$ holds (see [14]). This concept arises from the definition of *para-primal* algebras, introduced by D. M. Clark and P. H. Krauss [1], by simply omitting the requirement that $\mathfrak{A}$ be finite. Thus the finite locally para-primal algebras are exactly the para-primal algebras. It is easy to see that every locally quasi-primal algebra is locally para-primal.

1.3. Corollary. *An idempotent algebra* $\mathfrak{A} = (A; F)$ *is locally para-primal if and only if* $\mathfrak{A}^2$ *has no reduced subalgebras.*

Proof. The necessity is an immediate consequence of the definition of local para-primality. Conversely, if $\mathfrak{A}^2$ has no reduced subalgebras, then by Theorem 1.1 every reduced subuniverse B of any finite power of $\mathfrak{A}$ is as described in (1.1.2). Now it can be checked without difficulty that B satisfies the condition required in the definition of local para-primality. This implies that the same holds also for arbitrary subuniverses of finite powers of $\mathfrak{A}$. Hence $\mathfrak{A}$ is locally para-primal.

We note that for finite algebras $\mathfrak{A}$ Corollaries 1.2 and 1.3 can be strengthened further; see Corollaries 4.5 and 4.13 in [13].

Let $\mathscr{I}_A$ denote the clone of idempotent operations on A. Clearly, $\mathscr{I}_A$ is locally closed. Applying Theorem 1.1 we can determine the locally closed clones sitting "high up" in the lattice of locally closed subclones of $\mathscr{I}_A$ (in the terminology of [11] these clones, or more precisely the corresponding relations, form a generic system for $\mathscr{I}_A$, which is as irredundant as possible). We call the subsets

$$X^{a_1,a_2}=(A\times\{a_2\})\cup(\{a_1\}\times A)\quad (a_1,a_2\in A)$$

of A^2 *crosses*, and we write X^a for $X^{a,a}$ $(a\in A)$. For a subset B of some power of A the clone of all operations on A preserving B is denoted by $\mathrm{Pol}_A\{B\}$.

1.4. Corollary. *Let A be a set with $|A|\geqq 2$. Every locally closed, proper subclone of $\mathscr{I}_A$ is contained in one of the clones $\mathscr{I}_A\cap\mathrm{Pol}_A\{B\}$ where*

(1.4.1) $B\subset A$, $|B|\geqq 2$; *or*

$(1.4.2)_1$ $B=\pi_\square$ *for some permutation π of A with at most one fixed point such that all nontrivial cycles of π are of the same length q for some prime q; or*

$(1.4.2)_2$ $B=\pi_\square$ *for some permutation π of A with at most one fixed point such that all nontrivial cycles of π are of infinite length; or*

$(1.4.3)_1$ $B=X^a$ *for some $a\in A$; or*

$(1.4.3)_2$ $B=X^{a_1,a_2}$ *with $\{a_1,a_2\}=A$.*

These clones are locally closed, proper subclones of $\mathscr{I}_A$. The maximal locally closed subclones of $\mathscr{I}_A$ are exactly those of types (1.4.1), $(1.4.2)_1$, $(1.4.3)_1$ *and* $(1.4.3)_2$.

Proof. To prove the first claim let $\mathscr{C}$ be a proper subclone of $\mathscr{I}_A$, and assume $\mathscr{C}$ is locally closed. Then $\mathscr{C}=\mathscr{T}_{\mathrm{loc}}(\mathfrak{A})$ for the idempotent algebra $\mathfrak{A}=(A;\mathscr{C})$. If $\mathfrak{A}$ has a proper subuniverse B with $|B|\geqq 2$, then $\mathscr{C}\subseteqq\mathscr{I}_A\cap\mathrm{Pol}_A\{B\}$ with B of type (1.4.1), and we are done. Therefore we suppose from now on that the singletons and A are the only subuniverses of $\mathfrak{A}$. Since $\mathscr{C}\subset\mathscr{I}_A$, if $\mathfrak{A}$ is locally quasi-primal, then $\mathfrak{A}$ has a nonidentity automorphism σ. As the set of fixed points of each automorphism of $\mathfrak{A}$ is a subuniverse of $\mathfrak{A}$, it follows that every nonidentity power of σ has at most one fixed point. Thus, either $\pi=\sigma$ is of type $(1.4.2)_2$, or some power π of σ is of type $(1.4.2)_1$, implying in both cases that $\mathscr{C}\subseteqq\mathscr{I}_A\cap\mathrm{Pol}_A\{\pi_\square\}$. If $\mathfrak{A}$ is locally para-primal but not locally quasi-primal, then by Corollary 1.3 and Theorem 1.1 $\mathfrak{A}$ is locally term equivalent to the full idempotent reduct of the module $_{(\mathrm{End}\,_K\mathbf{A})}\mathbf{A}$ for some vector space $_K\mathbf{A}$ over a division ring K. Therefore every translation $x\pi=x+a$ with $a\neq 0$ is an automorphism of $\mathfrak{A}$ which is of type $(1.4.2)_1$ or $(1.4.2)_2$ according to whether the characteristic of K is prime or zero. Hence we conclude again that $\mathscr{C}\subseteqq\mathscr{I}_A\cap\mathrm{Pol}_A\{\pi_\square\}$. Finally, if $\mathfrak{A}$ is not locally para-primal, then by Corollary 1.3 $\mathfrak{A}^2$ has a reduced subuniverse B. Obviously, $B\lneqq A\times A$. Since the sets $B(a,x)$ and $B(x,a)$ are nonempty subuniverses of $\mathfrak{A}$ for all $a\in A$, it follows that B is a cross;

say $B=X^{a_1,a_2}$. Clearly, $X^{a_2,a_1}\approx X^{a_1,a_2}$, hence X^{a_2,a_1} is also a subuniverse of $\mathfrak{A}^2$. In case $a_1\neq a_2$ the set $\mathrm{pr}_1(X^{a_1,a_2}\cap X^{a_2,a_1})=\{a_1, a_2\}$ is a subuniverse of A, hence $a_1=a_2$ or $|A|=2$. Thus B is of type $(1.4.3)_1$ or $(1.4.3)_2$, and obviously $\mathscr{C}\subseteq\mathscr{I}_A\cap\cap\mathrm{Pol}_A\{B\}$. This concludes the proof of the first claim.

As regards the second assertion, it is straightforward to check that the clones $\mathscr{I}_A\cap\mathrm{Pol}_A\{B\}$ listed in the corollary are locally closed and are properly contained in $\mathscr{I}_A$. A case-by-case analysis shows also that, except for the obvious coincidences implied by the equalities

$$\mathrm{Pol}_A\{\pi_\square\}=\mathrm{Pol}_A\{(\pi^k)_\square\}\quad\text{if } \pi \text{ is as in } (1.4.2)_1,\quad 1<k<q,$$

and

$$\mathrm{Pol}_A\{X^{a_1,a_2}\}=\mathrm{Pol}_A\{X^{a_2,a_1}\}\quad(a_1, a_2\in A),$$

any two clones listed in the corollary and such that not both are of type $(1.4.2)_2$ are incomparable. Therefore the clones $\mathscr{I}_A\cap\mathrm{Pol}_A\{B\}$ with B of type (1.4.1), $(1.4.2)_1$, $(1.4.3)_1$ or $(1.4.3)_2$ are indeed maximal among the locally closed, proper subclones of $\mathscr{I}_A$. Finally, if $B=\pi_\square$ is of type $(1.4.2)_2$, then $\mathscr{I}_A\cap\mathrm{Pol}_A\{\pi_\square\}$ is not maximal, since

$$\mathscr{I}_A\cap\mathrm{Pol}_A\{\pi_\square\}\subset\mathscr{I}_A\cap\mathrm{Pol}_A\{(\pi^l)_\square\}\quad(\subset\mathscr{I}_A)\quad\text{for every integer}\quad l>1.$$

The proof of the corollary is complete.

We now sketch the proof of Theorem 1.1. The first lemma is the same as Lemma 4.4 in [13], excepting that the algebra is not assumed to be finite. As the proof carries over without change to this more general situation, we do not go into the details here.

1.5. Lemma. *Let $\mathfrak{A}=(A; F)$ be an idempotent algebra with $|A|>1$, and assume $\mathfrak{A}^2$ has no reduced subuniverses of size m for some cardinal m $(1<m\leqq|A|)$. Furthermore, let $B\preceq B_1\times\ldots\times B_n$ $(n\geqq 2)$ be a directly indecomposable subuniverse of $\mathfrak{A}^n$ of size m. Then*

(1.5.1) *$\mathfrak{B}_i=(B_i; F)$ $(1\leqq i\leqq n)$ are isomorphic subalgebras of $\mathfrak{A}$, and*

(1.5.2) *for arbitrary isomorphisms π_i: $\mathfrak{B}_i\to\mathfrak{B}_1$ $(1\leqq i\leqq n)$ the subuniverse $B[\pi_1, \ldots, \pi_n]$ of $\mathfrak{A}^n$ has the form* (1) *for some $u\in\mathbf{n}$ $(u\geqq 2)$ and for some operations $g_u, \ldots, g_n\in\mathscr{O}_{B_1}^{(u-1)}$.*

The following statement is an important intermediate step in the proof of Lemma 1.5 (cf. Claim 1 in the proof of Lemma 4.4 in [13]).

1.6. Lemma. *Under the assumptions of Lemma 1.5 B has a projection*

$$\bar{B}=\mathrm{pr}_{i_1,\ldots,i_k}B\preceq B_{i_1}\times\cdots\times B_{i_k}\quad(\{i_1, \ldots, i_k\}\subseteq\mathbf{n})$$

with $k\geqq 2$ such that

(1.6.1) $|B_{i_l}|=m$ *for all* l $(1\leqq l\leqq k)$.

(1.6.2) $\mathrm{pr}_{\mathbf{k}-\{j\}}\bar{B}=\prod_{l\in\mathbf{k}-\{j\}} B_{i_l}$ *for all* j $(1\leqq j\leqq k)$, *and*

(1.6.3) $\bar{B}(x_1, b_2, \ldots, b_{j-1}, x_2, b_{j+1}, \ldots, b_k)$ *is a bijection* $B_{i_1}\to B_{i_j}$ *for all* j $(2\leqq j\leqq k)$ *and for all elements* $b_l\in B_{i_l}$ $(2\leqq l\leqq k,\ l\neq j)$.

Notice that condition (1.6.3) implies the existence of a function $g\colon B_{i_2}\times\ldots\times\times B_{i_k}\to B_{i_1}$ such that $\bar{B}=g_{\square}$, moreover, $g(b_2, \ldots, b_{j-1}, x, b_{j+1}, \ldots, b_k)\colon B_{i_j}\to B_{i_1}$ is a bijection for all j $(2\leqq j\leqq k)$ and for all elements $b_l\in B_{i_l}$ $(2\leqq l\leqq k,\ l\neq j)$. This leads to the following definition. A function $h\colon C_2\times\ldots\times C_m\to C_1$ $(m\geqq 2, C_1, \ldots, C_m\subseteq\subseteq A)$ is said to have the *constant substitution property* iff for every j $(2\leqq j\leqq m)$ such that h depends on its j-th variable, and for arbitrary elements $c_l\in C_l$ $(2\leqq l\leqq m,\ l\neq j)$, the unary function

$$h(c_2, \ldots, c_{j-1}, x, c_{j+1}, \ldots, c_m)\colon C_j\to C_1$$

is a bijection.

The next result is a special case of Theorem 2.1 in [14] (cf. also Proposition 3.4 in [13]).

1.7. Proposition. *Let* B *be a set with* $|B|>1$, *and* $\mathscr{C}$ *a clone on* B *containing all the constants. Assume*

(1.7.1) *every surjective operation in* $\mathscr{C}$ *has the constant substitution property,*

(1.7.2) $\mathscr{C}$ *contains a surjective operation depending on at least two of its variables, and*

(1.7.3) *for every quasigroup operation in* $\mathscr{C}$, $\mathscr{C}$ *also contains the corresponding left and right divisions.*

Then there exist a division ring K *and a vector space* ${}_K\mathbf{B}=(B;\,+,K)$ *such that* $\mathscr{C}=\mathscr{P}({}_K\mathbf{B})$.

Now we are in a position to prove the theorem.

Proof of Theorem 1.1. Let B be of size m, and assume (1.1.1) fails for $\mathfrak{A}$. Then $n>2$ and the conclusions of Lemma 1.5 hold for B. Therefore we have to prove only the claims for $\mathfrak{B}_1$ and that in the representation (1) of $B[\pi_1, \ldots, \pi_n]$ we have $g_u, \ldots, g_n\in\mathscr{P}({}_K\mathbf{B}_1)$.

In what follows, all operations occurring are defined on B_1. Let $\mathscr{C}$ denote the set of all operations commuting with every basic operation of $\mathfrak{B}_1$ (and hence with every local term operation of $\mathfrak{B}_1$). It is easy to see that $\mathscr{C}$ is a clone on B_1 satisfying (1.7.3). Since $\mathfrak{B}_1$ is idempotent, $\mathscr{C}$ contains all the constants. Furthermore, every operation g_j $(u\leqq j\leqq n)$ occurring in the representation (1) of $B[\pi_1, \ldots, \pi_n]$ described in Lemma 1.5 belongs to $\mathscr{C}$, as $(g_j)_{\square}^{\cdot}$ is a projection of $B[\pi_1, \ldots, \pi_n]$. These operations g_j $(u\leqq j\leqq n)$ are obviously surjective; moreover, since (1.1.1) fails and B is reduced, therefore each of them depends on at least two of its variables. Finally, we show that every surjective operation $g\in\mathscr{C}$ has the constant substitution property. We may assume without loss of generality that g depends on all of its variables. If, say, g is k-ary, then $g_\Box$ is a directly indecomposable subuniverse of $\mathfrak{A}^{k+1}$ of size m. As no proper projection of $g_\Box$ can satisfy condition (1.6.3), we conclude that (1.6.3) holds for $g_\Box$, implying that g has the constant substitution property.

Thus Proposition 1.7 applies for the clone $\mathscr{C}$. Consequently there exists a vector space ${}_K\mathbf{B}_1=(B_1; +, K)$ over some division ring K such that $\mathscr{C}=\mathscr{P}({}_K\mathbf{B}_1)$. Hence $g_u, \dots, g_n\in\mathscr{P}({}_K\mathbf{B}_1)$. Furthermore, it is easy to see that the clone of the full idempotent reduct of the module ${}_{(\mathrm{End}\,{}_K\mathbf{B}_1)}\mathbf{B}_1$ coincides with the clone $\mathscr{C}^*$ of all operations commuting with each member of $\mathscr{C}$. Therefore it remains to prove that $\mathscr{T}_{\mathrm{loc}}(\mathfrak{B}_1)=\mathscr{C}^*$. The inclusion $\subseteq$ is trivial by the definition of $\mathscr{C}$.

Before verifying the reverse inclusion observe that the singletons are the only proper subuniverses of $\mathfrak{B}_1$. Indeed, if $S\subset B_1$ $(S\neq\emptyset)$ is a proper subuniverse of $\mathfrak{B}_1$, then $x_1-x_2\in\mathscr{C}$ implies that

$$U=\{(x_1,x_2)\in B_1^2\colon\ x_1-x_2\in S\}$$

is a subuniverse of $\mathfrak{B}_1^2$ (and hence of $\mathfrak{A}^2$). Since $\mathrm{pr}_1 U=\mathrm{pr}_2 U=B_1$, $U\neq B_1^2$, and by assumption U is not reduced, therefore it follows that U is a bijection. Hence $|S|=1$.

Now let $f\in\mathscr{C}^*$, and let C be an arbitrary directly indecomposable subuniverse of $\mathfrak{B}_1^t$ for some integer $t\geqq 1$. Since $\mathfrak{B}_1$ has no nonsingleton proper subalgebras, we have either $C\leqq B_1^t$ so that C is of size m, or $|C|=t=1$. If $t=1$, then f obviously preserves C. Suppose $t\geqq 2$. Lemma 1.5 implies then that C has the form

$$C\approx\{(y_1,\dots,y_{v-1},f_v(y_1,\dots,y_{v-1}),\dots,f_t(y_1,\dots,y_{v-1}))\colon\ y_1,\dots,y_{v-1}\in B_1\}$$

for some v $(2\leqq v\leqq t)$ and some operations $f_v,\dots,f_t\in\mathcal{O}_{B_1}$. The sets $(f_j)_\Box$ $(v\leqq j\leqq t)$ are projections of C, yielding that $f_v,\dots,f_t\in\mathscr{C}$. Thus f commutes with $f_v,\dots,f_t$, implying that f preserves C as well. This means that f preserves every directly indecomposable subuniverse of each finite power of $\mathfrak{B}_1$. Hence it preserves also all subuniverses of finite powers of $\mathfrak{B}_1$, that is, $f\in\mathscr{T}_{\mathrm{loc}}(\mathfrak{B}_1)$. Therefore $\mathscr{C}^*=$ $=\mathscr{T}_{\mathrm{loc}}(\mathfrak{B}_1)$, as was to be proved.

2. Plain idempotent algebras

Recall that an algebra is called *plain* iff it is simple and has no nonsingleton proper subalgebras. Clearly, for an idempotent algebra the property of having no nonsingleton proper subalgebras implies simplicity. As we shall see in this section, having no nonsingleton proper subalgebras is a rather strong constraint on idempotent algebras: up to local term equivalence, there are only "a few" plain idempotent algebras.

In the description of plain idempotent algebras an important role will be played by the "higher dimensional crosses"

$$X_n^a = \bigcup_{i=1}^{n} (A\times\ldots\times A\times\overset{i}{\{a\}}\times A\times\ldots\times A), \quad n \geqq 2,$$

where a is a fixed element of the set A. Obviously, $X_2^a = X^a$. For $k \geqq 2$ and $a \in A$ let $\mathcal{F}_k^a$ denote the clone of all idempotent operations on A preserving X_k^a. Furthermore, put $\mathcal{F}_\omega^a = \bigcap_{2 \leqq k < \omega} \mathcal{F}_k^a$. Since for arbitrary element $b \in A$, $b \neq a$, we have $X_n^a(x_1, \ldots, x_{n-1}, b) = X_{n-1}^a$, therefore

$$\mathcal{F}_2^a \supseteqq \mathcal{F}_3^a \supseteqq \ldots \supseteqq \mathcal{F}_k^a \supseteqq \mathcal{F}_{k+1}^a \supseteqq \ldots \supseteqq \mathcal{F}_\omega^a.$$

For a permutation group G acting on A we will denote by $\mathcal{I}_A(G)$ the clone of all idempotent operations on A commuting with every member of G.

2.1. Theorem. *Every plain idempotent algebra* $\mathfrak{A} = (A; F)$ *with* $|A| \geqq 3$ *is locally term equivalent to one of the following algebras:*

(2.1.1) $(A; \mathcal{I}_A(G))$ *for a permutation group* G *acting on* A *such that every nonidentity member of* G *has at most one fixed point;*

(2.1.2) *the full idempotent reduct of the module* $_{(\mathrm{End}\,_K\mathbf{A})}\mathbf{A}$ *for some vector space* $_K\mathbf{A} = (A; +, K)$ *over a division ring* K;

(2.1.3) $(A; \mathcal{I}_A(G) \cap \mathcal{F}_k^0)$ *for some* k $(2 \leqq k \leqq \omega)$, *some element* $0 \in A$, *and a permutation group* G *acting on* A *such that* 0 *is the unique fixed point of every nonidentity member of* G.

Remarks. 1. It is not hard to show that every algebra locally term equivalent to an algebra in (2.1.1) or (2.1.3) is plain. The same is well known to hold for (2.1.2), too. Note that the algebras in (2.1.1) are locally quasi-primal.

2. The conclusion of the theorem fails for 2-element algebras. Obviously, every 2-element algebra is plain, and Post's description [8] of all clones on a 2-element set (or Corollary 1.4 above) shows that, up to term equivalence, there are more 2-element idempotent algebras than those of types (2.1.1)—(2.1.3) listed in the theorem.

The first step of the proof of Theorem 2.1 is based on Theorem 1.1.

2.2. Proposition. *For a plain idempotent algebra* $\mathfrak{A} = (A; F)$ *with* $|A| \geqq 3$ *one of the following conditions holds:*

(2.2.1) $\mathfrak{A}$ *is locally quasi-primal, or*

(2.2.2) *there exist a division ring* K *and a vector space* $_K\mathbf{A} = (A; +, K)$ *such that* $\mathfrak{A}$ *is locally term equivalent to the full idempotent reduct of the module* $_{(\mathrm{End}\,_K\mathbf{A})}\mathbf{A}$, *or*

(2.2.3) *there exists an element* $0 \in A$ *such that* X^0 *is the only reduced subuniverse of* $\mathfrak{A}^2$, *moreover,* 0 *is the unique fixed point of each nonidentity automorphism of* $\mathfrak{A}$.

Proof. Since $\mathfrak{A}$ has no nonsingleton proper subalgebras, Theorem 1.1 implies that $\mathfrak{A}$ is of type (2.2.1) or (2.2.2), or $\mathfrak{A}^2$ has a reduced subuniverse. Assume the last possibility holds for $\mathfrak{A}$, and let D be a reduced subuniverse of $\mathfrak{A}^2$. Obviously, $D \leqq A \times A$. Using that the subuniverses $D(x, a)$, $D(a, x)$ $(a \in A)$ of $\mathfrak{A}$ are singletons or equal to A, we can conclude that D is a cross, say $D = X^{a_1, a_2}$ $(a_1, a_2 \in A)$. We must have $a_1 = a_2$, since otherwise $\mathrm{pr}_1(X^{a_1, a_2} \cap X^{a_2, a_1}) = \{a_1, a_2\}$ would be a nonsingleton proper subuniverse of $\mathfrak{A}$. It follows now that there is at most one cross among the subuniverses of $\mathfrak{A}^2$. Indeed, if X^b and X^c $(b, c \in A, b \neq c)$ were subuniverses of $\mathfrak{A}^2$, then $\mathrm{pr}_1(X^b \cap X^c) = \{b, c\}$ would again be a nonsingleton proper subuniverse of $\mathfrak{A}$. Thus there is an element $0 \in A$ such that X^0 is the unique reduced subuniverse of $\mathfrak{A}^2$. This implies in particular that $X^0[\pi, \pi] = X^0$ for arbitrary automorphism π of $\mathfrak{A}$; hence π fixes 0. Furthermore, since the fixed points of an automorphism of $\mathfrak{A}$ form a subuniverse in $\mathfrak{A}$, 0 is the only fixed point of each nonidentity automorphism of $\mathfrak{A}$.

Now we discuss in more detail the algebras $\mathfrak{A}$ of type (2.2.3). To show that every such $\mathfrak{A}$ is locally term equivalent to an algebra in (2.1.3), we determine the subuniverses of finite powers of $\mathfrak{A}$. For a natural number $n \geqq 2$ and for a family P of subsets of $\mathbf{n}$ we set

$$Y^0_{n,P} = \bigcup_{I \in P} A^{(n,I)}$$

where

$$A^{(n,I)} = A_1 \times \ldots \times A_n \quad \text{with} \quad A_i = \begin{cases} A & \text{if } i \in I \\ \{0\} & \text{if } i \in \mathbf{n} - I. \end{cases}$$

Since the element 0 is fixed throughout this discussion, we omit the superscript ^{0} in $Y^0_{n,P}$ and X^0_n. Clearly, $Y_{n,P} \subseteqq X_n$ unless $\mathbf{n} \in P$, and equality holds if P is the set of $(n-1)$-element subsets of $\mathbf{n}$. Let us call a subset C of A^n *irredundant* iff $\mathrm{pr}_i C = A$ for all $i \in \mathbf{n}$ and no projection $\mathrm{pr}_{i,j} C$ $(i, j \in \mathbf{n},\ i \neq j)$ of C is a permutation of A. Clearly, all reduced subsets of A^n are irredundant.

2.3. Proposition. *Let* $\mathfrak{A} = (A; F)$ $(|A| \geqq 3)$ *be a plain idempotent algebra satisfying condition* (2.2.3). *Then for every integer* $n \geqq 2$, *every irredundant subuniverse of* $\mathfrak{A}^n$ *is of the form* $Y_{n,P}$ *for some family* P *of subsets of* $\mathbf{n}$.

The case $n=2$ is a consequence of (2.2.3): A^2 and X_2 are the only irredundant subuniverses of $\mathfrak{A}^2$. The next three steps of the proof will be carried out in Lemmas 2.4 through 2.6.

2.4. Lemma. *The claim of Proposition* 2.3 *is true for* $n=3$.

Proof. Let C be an irredundant subuniverse of $\mathfrak{A}^3$. Clearly, its projections $\mathrm{pr}_{i,j} C$ $(i, j \in 3, i \neq j)$ are also irredundant, and hence equal A^2 or X_2.

If $\mathrm{pr}_{1,2}C=\mathrm{pr}_{1,3}C=\mathrm{pr}_{2,3}C=X_2$, then every triple from C has at least two components 0. Thus

$$C \subseteq (A\times\{0\}\times\{0\})\cup(\{0\}\times A\times\{0\})\cup(\{0\}\times\{0\}\times A). \tag{2}$$

Since $\mathrm{pr}_1 C=A$, the subuniverse $C(x_1, 0, 0)$ of $\mathfrak{A}$ must contain $A-\{0\}$. Thus the assumption $|A|\geqq 3$ and the plainness of $\mathfrak{A}$ imply that $C(x_1, 0, 0)=A$, hence $A\times\{0\}\times\{0\}\subseteq C$. By symmetry it follows that equality holds in (2).

Suppose $\mathrm{pr}_{1,2}C=\mathrm{pr}_{1,3}C=X_2$, $\mathrm{pr}_{2,3}C=A^2$. Then, clearly,

$$C \subseteq (A\times\{0\}\times\{0\})\cup(\{0\}\times A\times A). \tag{3}$$

Since for arbitrary $(a, b)\in A^2$ with $a\neq 0$ we have $(a, b)\in \mathrm{pr}_{2,3}C$, therefore $(x, a, b)\in C$ for some $x\in A$. However, $\mathrm{pr}_{1,2}C=X_2$ yields that $x=0$. Hence for all $b\in A$ the subuniverse $C(0, x_1, b)$ contains the set $A-\{0\}$, implying that $C(0, x_1, b)=A$. Thus $\{0\}\times A\times A\subseteq C$, which together with $\mathrm{pr}_1 C=A$ shows that we have equality in (3).

Assume now that $\mathrm{pr}_{1,2}C=X_2$ and $\mathrm{pr}_{1,3}C=\mathrm{pr}_{2,3}C=A^2$. Then

$$C \subseteq X_2\times A = (A\times\{0\}\times A)\cup(\{0\}\times A\times A).$$

As in the previous case, we get that $\{0\}\times A\times A\subseteq C$, and similarly (interchanging the role of the first and second components) $A\times\{0\}\times A\subseteq C$. Thus $C=X_2\times A$.

By symmetry it remains to consider the case $\mathrm{pr}_{1,2}C=\mathrm{pr}_{1,3}C=\mathrm{pr}_{2,3}C=A^2$. If $C=A^3$, we are done, so assume that $C\neq A^3$. First we show the required equality $C=X_3$ under the additional assumption that there is an element $c\in A$ with $C(c, x_1, x_2)=A^2$. In this case we have $C(x_1, x_2, b)\supseteq\{c\}\times A$ for all $b\in A$. Taking into account that $C(x_1, x_2, b)$ is a subuniverse of $\mathfrak{A}^2$ and $\mathrm{pr}_i C(x_1, x_2, b)=A$ for $i\in\mathbf{2}$ (the latter follows from $\mathrm{pr}_{1,3}C=\mathrm{pr}_{2,3}C=A^2$), we get that $C(x_1, x_2, b)$ equals X_2 or A^2 for every $b\in A$. Since $C\neq A^3$, the former has to hold for at least one $b\in A$, implying that $c=0$. On the other hand, $\mathrm{pr}_{1,2}C=A^2$ ensures that there is a $b'\in A$ with $C(x_1, x_2, b')=A^2$. Therefore the same argument as before (with c replaced by b') yields that $C(x_1, x_2, b)=A^2$ if and only if $b=0$. Hence

$$C = \bigcup_{b\in A}(C(x_1, x_2, b)\times\{b\}) = (A^2\times\{0\})\cup\bigcup_{\substack{b\in A\\ b\neq 0}}(X_2\times\{b\}) = X_3.$$

Finally, suppose that for all elements $c\in A$ the subuniverses $C(c, x_1, x_2)$, and symmetrically also $C(x_1, c, x_2)$, $C(x_1, x_2, c)$, of $\mathfrak{A}^2$ are distinct from A^2. Since their projections onto each component are equal to A, we get that each of these subuniverses is either X_2 or an automorphism of $\mathfrak{A}$. Since all automorphisms of $\mathfrak{A}$ fix 0, and $\mathrm{pr}_{1,2}C=A^2$, we conclude that there exists an element $b\in A$ such that $C(x_1, x_2, b)=X_2$. Then $C(0, x_1, x_2)\supseteq A\times\{b\}$, implying that $C(0, x_1, x_2)=X_2$ and

$b=0$. This shows that $C(x_1, x_2, c)=X_2$ if and only if $c=0$, and by symmetry the same holds for the subuniverses $C(x_1, c, x_2)$, $C(c, x_1, x_2)$ as well.

Consider now the set

$$D = \{(x, y, z)\in A^3\colon \text{ there is a } u\in A \text{ such that } (x, y, u),\ (u, y, z)\in C\}.$$

It is easy to check that D is a subuniverse of $\mathfrak{A}^3$. Furthermore, $\mathrm{pr}_{1,2}C=\mathrm{pr}_{2,3}C=A^2$ implies that $\mathrm{pr}_{1,2}D=\mathrm{pr}_{2,3}D=A^2$. Choosing $u=0$ in the definition of D we get that $A\times\{0\}\times A\subseteq D$. Hence D satisfies the additional assumption under which we can conclude that D equals X_3 or A^3. Then for arbitrary elements $a, b\in A-\{0\}$ we have $(0, a, b)\in D$, that is $(0, a, c), (c, a, b)\in C$ for some $c\in A$. However, $C(0, x_1, x_2)=X_2$ implies that $c=0$ and $(a, b)\in X_2$, a contradiction. Thus this case cannot occur, completing the proof of Lemma 2.4.

2.5. Lemma. *For arbitrary integer $n\geqq 3$ and for every subuniverse C of $\mathfrak{A}^n$ satisfying $\mathrm{pr}_{\mathbf{n}-\{i\}}C=A^{n-1}$ for all $i\in\mathbf{n}$, we have $C=A^n$ or $C=X_n$.*

Proof. We proceed by induction on n. Clearly C is irredundant, therefore by Lemma 2.4 the claim is true if $n=3$. Let now $n\geqq 4$, and suppose $C\neq A^n$. Since the subuniverses $C(x_1, \ldots, x_{i-1}, a, x_i, \ldots, x_{n-1})$ $(i\in\mathbf{n},\ a\in A)$ of $\mathfrak{A}^{n-1}$ satisfy the assumption of the lemma, we get from the induction hypothesis that

$$C(x_1, \ldots, x_{i-1}, a, x_i, \ldots, x_{n-1})=A^{n-1} \text{ or } X_{n-1} \text{ for all } a\in A \text{ and } i\in\mathbf{n}.$$

We have $C(b, x_1, \ldots, x_{n-1})=A^{n-1}$ for at least one $b\in A$, because $\mathrm{pr}_{\mathbf{n}-1}C=A^{n-1}$. Since $C\neq A^n$, there also exists an element $c\in A$ such that $C(x_1, \ldots, x_{n-1}, c)=X_{n-1}$. Thus

$$X_{n-1} = C(x_1, \ldots, x_{n-1}, c) \supseteq \{b\}\times A^{n-2},$$

yielding $b=0$. Consequently

$$C(b, x_1, \ldots, x_{n-1}) = \begin{cases} A^{n-1} & \text{if } b=0 \\ X_{n-1} & \text{otherwise} \end{cases} \quad (b\in A),$$

whence $C=X_n$, as required.

2.6. Lemma. *For arbitrary irredundant subuniverse C of $\mathfrak{A}^n$ $(n\geqq 3)$ the subuniverses $C(x_1, \ldots, x_{i-1}, 0, x_i, \ldots, x_{n-1})$ $(i\in\mathbf{n})$ of $\mathfrak{A}^{n-1}$ are also irredundant.*

Proof. By symmetry it suffices to prove the statement for $\tilde{C}=C(x_1, \ldots, x_{n-1}, 0)$. Firstly, let $1\leqq i\leqq n-1$. As $D=\mathrm{pr}_{i,n}C$ is an irredundant subuniverse of $\mathfrak{A}^2$, we must have $D=A^2$ or $D=X_2$. Thus $\mathrm{pr}_i\tilde{C}=D(x, 0)=A$. Secondly, let $1\leqq i<j\leqq n-1$. Clearly, $D'=\mathrm{pr}_{i,j,n}C$ is an irredundant subuniverse of $\mathfrak{A}^3$. Thus, making use of Lemma 2.4, one can easily see that the set $\mathrm{pr}_{i,j}\tilde{C}=D'(x_1, x_2, 0)$ is not a permutation of A. This shows that $\tilde{C}$ is irredundant.

Proof of Proposition 2.3. The case $n=3$ is settled in Lemma 2.4, so we may assume that $n\geqq 4$ and the claim is true for the irredundant subuniverses of $\mathfrak{A}^{n-1}$. Consider an irredundant subuniverse C of $\mathfrak{A}^n$. If C contains an n-tuple with all components distinct from 0, then by the induction hypothesis $\mathrm{pr}_{\mathbf{n}-\{i\}}C=A^{n-1}$ for all $i\in\mathbf{n}$. Hence, by Lemma 2.5, $C=A^n$. Otherwise, if every n-tuple in C has at least one component 0, then Lemma 2.6 and the induction hypothesis ensure for each $i\in\mathbf{n}$ the existence of a family P_i of subsets of $\mathbf{n}-\{i\}$ such that

$$C(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) = Y_{n-1, P_i}.$$

Putting $P=\bigcup_{i\in\mathbf{n}} P_i$ we get that $C=Y_{n,P}$.

Making use of Proposition 2.3 we can now conclude the proof of Theorem 2.1.

Proof of Theorem 2.1 Let G denote the automorphism group of $\mathfrak{A}$. According to Proposition 2.2 we have to distinguish three cases. Suppose first (2.2.1), that is, $\mathfrak{A}$ is locally quasi-primal. Since $\mathfrak{A}$ is plain, every internal isomorphism of $\mathfrak{A}$ is either an automorphism of $\mathfrak{A}$, or an isomorphism between two singleton subalgebras of $\mathfrak{A}$. Therefore $\mathscr{T}_{\mathrm{loc}}(\mathfrak{A})=\mathscr{I}_A(G)$. Furthermore, since the set of fixed points of each automorphism of $\mathfrak{A}$ is a subuniverse of $\mathfrak{A}$, therefore each nonidentity member of G has at most one fixed point. Consequently $\mathfrak{A}$ is locally term equivalent to an algebra of type (2.1.1).

In case (2.2.2) we have nothing to prove.

Finally, if (2.2.3) holds for $\mathfrak{A}$, then we can apply Proposition 2.3. Observe first that for arbitrary irredundant subuniverse C of $\mathfrak{A}^n$ $(n\geqq 3)$ such that $C\subset X_n$ we have

$$C = \{(x_1, \ldots, x_n)\in A^n\colon (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)\in \mathrm{pr}_{\mathbf{n}-\{i\}}C \text{ for all } i\in\mathbf{n}\}.$$

Indeed, the inclusion $\subseteqq$ being trivial, suppose $(x_1, \ldots, x_n)\in A^n$ belongs to the set on the right hand side. Since $C\subset X_n$, by Lemma 2.5 some projection of C onto $n-1$ components is distinct from A^{n-1}, say $\mathrm{pr}_{\mathbf{n}-1}C\neq A^{n-1}$. However, $\mathrm{pr}_{\mathbf{n}-1}C$ is an irredundant subuniverse of $\mathfrak{A}^{n-1}$, therefore by Proposition 2.3 at least one component of $(x_2, \ldots, x_n)\in\mathrm{pr}_{\mathbf{n}-1}C$ equals 0, say $x_2=0$. Then $(x_1, x_3, \ldots, x_n)\in\mathrm{pr}_{\mathbf{n}-\{2\}}C$ implies that $(x_1, y, x_3, \ldots, x_n)\in C$ for some $y\in A$. Since C is of the form described in Proposition 2.3, we have also $(x_1, x_2, x_3, \ldots, x_n)=(x_1, 0, x_3, \ldots, x_n)\in C$.

Since $\mathfrak{A}$ is a plain idempotent algebra, an idempotent operation on A is a local term operation of $\mathfrak{A}$ if and only if it preserves the automorphisms of $\mathfrak{A}$ and the irredundant subuniverses of finite powers of $\mathfrak{A}$. By repeated application of the observation made in the previous paragraph it follows that for every irredundant subuniverse C of $\mathfrak{A}^n$ $(n\geqq 1)$, an operation preserves C if and only if it preserves all those projections of C which are of the form X_j for some j $(2\leqq j\leqq n)$. Thus $\mathscr{T}_{\mathrm{loc}}(\mathfrak{A})=\mathscr{I}_A(G)\cap\mathscr{F}_\omega^0$ if all X_j $(j\geqq 2)$ occur among the subuniverses of finite powers of $\mathfrak{A}$, and $\mathscr{T}_{\mathrm{loc}}(\mathfrak{A})=\mathscr{I}_A(G)\cap\mathscr{F}_k^0$ $(2\leqq k<\omega)$ if k is the largest j such that X_j is a subuniverse of $\mathfrak{A}^j$. The proof of Theorem 2.1 is complete.

Recall that an algebra $\mathfrak{A}=(A;\, F)$ is *locally functionally complete* iff $\mathscr{P}_{\mathrm{loc}}(\mathfrak{A})=\mathscr{O}_A$. In terms of subuniverses this condition is equivalent to requiring that for every integer $n\geqq 1$, A^n is the only irredundant reflexive subuniverse of $\mathfrak{A}^n$. ($B\subseteqq A^n$ is said to be *reflexive* iff $(a, \ldots, a)\in B$ for all $a\in A$.) Thus a slight improvement of a result of L. Szabó [12] can easily be derived from Propositions 2.2 and 2.3.

2.7. Corollary. *If $\mathfrak{A}=(A;\, F)$ is a plain idempotent algebra with $|A|\geqq 3$, then either $\mathfrak{A}$ is locally functionally complete, or there exist a division ring K and a vector space ${}_K\mathbf{A}=(A;\, +, K)$ such that $\mathfrak{A}$ is locally term equivalent to the full idempotent reduct of the module ${}_{(\mathrm{End}\,_K\mathbf{A})}\mathbf{A}$.*

Proof. By Proposition 2.2 we have to show that if $\mathfrak{A}$ is of type (2.2.1) or (2.2.3), then for every integer $n\geqq 1$, A^n is the only irredundant reflexive subuniverse of $\mathfrak{A}^n$. For type (2.2.1) this is well known (see P. H. Krauss [2]), while for type (2.2.3) it follows from Proposition 2.3.

3. Two results on minimal clones

Throughout this section projections will be called *trivial operations,* and the term *trivial clone* will mean the clone of projections. It is obvious that a minimal clone is generated by each of its nontrivial members. Thus the most natural way of classifying minimal clones is by their nontrivial members of least possible arity. Accordingly, by a result of I. G. Rosenberg [10], the algebras $(A;\, f)$ with minimal clones fall into five types: (i) f is a nontrivial unary operation, (ii) f is a nontrivial idempotent binary operation, (iii) f is a majority operation, (iv) f is a nontrivial semiprojection, or (v) $f(x, y, z)=x+y+z$ for an elementary Abelian 2-group.

It is well known (see J. Płonka [7]) that for every elementary Abelian q-group $\mathbf{A}=(A;\, +)$ (q prime) the algebra $(A;\, x-y+z)$ has a minimal clone. If $q>2$, then these algebras are of type (ii), since they have nontrivial binary term operations. In this section we present two conditions ensuring that a finite algebra with minimal clone be term equivalent to an algebra of this form.

As was observed by I. G. Rosenberg [10], an algebra $(A;\, f)$ where f is a minority operation has a minimal clone if and only if $f(x, y, z)=x+y+z$ for some elementary Abelian 2-group $\mathbf{A}=(A;\, +)$. P. P. Pálfy posed a more general question: For which Mal'tsev operations p is the clone of the algebra $(A;\, p)$ minimal? The following result answers this question in the finite case.

3.1. Theorem. *A finite algebra $(A; p)$ where p is a Mal'tsev operation has a minimal clone if and only if there exists an elementary Abelian group $\mathbf{A}=(A; +)$ such that $p(x, y, z)=x-y+z$.*

It remains open whether the same holds true for infinite algebras $(A; p)$ as well.

The next result gives a characterization for those idempotent groupoids with minimal clones which are term equivalent to $(A; x-y+z)$.

3.2. Theorem. *A finite idempotent groupoid $(A; \cdot)$ with minimal clone is term equivalent to an algebra $(A; x-y+z)$ for some elementary Abelian q-group (q is an odd prime) if and only if it has a minimal nonsingleton subgroupoid of cardinality greater than 2.*

The if part of this statement can be rephrased as follows: In a finite idempotent groupoid $(A; \cdot)$ with minimal clone every minimal nonsingleton subgroupoid is 2-element, unless $(A; \cdot)$ is term equivalent to $(A; x-y+z)$ for some elementary Abelian q-group (q is an odd prime). Since the clone of every subgroupoid of $(A; \cdot)$ is minimal or trivial, it follows that every 2-element subgroupoid of $(A; \cdot)$ is either a left zero semigroup, or a right zero semigroup, or a semilattice. This suggests that in trying to determine the finite idempotent groupoids with minimal clones it may be useful to classify these groupoids according to the types of their 2-element subgroupoids. P. P. Pálfy [4] has made an interesting observation in this direction by proving that if an idempotent groupoid $(A; \cdot)$ with minimal clone has a left zero semigroup as well as a right zero semigroup among its 2-element subgroupoids, then $(A; \cdot)$ is a rectangular band.

We now turn to the proof of Theorems 3.1 and 3.2. As a preparation, we state two lemmas.

3.3. Lemma. *A finite plain idempotent algebra $\mathfrak{B}=(B; F)$ with $|B|\geqq 3$ has a minimal clone if and only if it is term equivalent to $(B; x-y+z)$ for some cyclic group $\mathbf{B}=(B; +)$ of prime order.*

Proof. Let $\mathfrak{B}=(B; F)$ be a finite plain idempotent algebra with $|B|\geqq 3$. By Theorem 2.1 we have one of the following three possibilities for $\mathfrak{B}$:

(a) $\mathfrak{B}$ is quasi-primal, or

(b) there exist a prime q and an elementary Abelian q-group $\mathbf{B}=(B; +)$ such that $\mathfrak{B}$ is affine with respect to $\mathbf{B}$, or

(c) $\mathcal{I}_A(G)\cap\mathcal{F}_\omega^0\subseteq\mathcal{T}(\mathfrak{B})$ for some element $0\in B$ and some permutation group G acting on B such that 0 is the unique fixed point of each nonidentity permutation in G.

Suppose $\mathfrak{B}$ has a minimal clone. Since the clone of a quasi-primal algebra cannot be minimal, case (a) does not hold for $\mathfrak{B}$. We prove that case (c) is also excluded. Assume $\mathfrak{B}$ satisfies the conditions in (c), and define a binary operation $*$ on B as follows:

$$x*y=\begin{cases}0 & \text{if} \quad x=0 \quad \text{or} \quad y=0 \\ x & \text{otherwise}\end{cases} \qquad (x, y\in B).$$

It is easy to check that $*\in\mathcal{I}_A(G)\cap\mathcal{F}_\omega^0$. Hence $*$ is a nontrivial term operation of $\mathfrak{B}$. However, the algebra $(B; *)$ is not term equivalent to $\mathfrak{B}$, because it is not plain (every 2-element subset is a subalgebra). Therefore the clone of $\mathfrak{B}$ is not minimal. Thus $\mathfrak{B}$ satisfies condition (b), implying that $x-y+z$ is a term operation of $\mathfrak{B}$. Since $\mathcal{T}(\mathfrak{B})$ is a minimal clone and $\mathfrak{B}$ is plain, we conclude that $\mathfrak{B}$ is term equivalent to $(B; x-y+z)$ and $|B|=q$. This completes the proof of the only if part. The converse is well known (cf. J. Płonka [7]).

3.4. Lemma. *Let $\mathfrak{A}=(A; f)$ be a finite idempotent algebra whose clone is minimal, and let $\mathfrak{B}=(B; f)$ be a subalgebra of $\mathfrak{A}$. If $\mathfrak{B}$ is term equivalent to $(B; x-y+z)$ for some cyclic group $\mathbf{B}=(B; +)$ of prime order, then there exists an elementary Abelian group $\mathbf{A}=(A; +)$ such that $\mathfrak{A}$ is term equivalent to $(A; x-y+z)$.*

Proof. Let $|B|=q$ (q prime). For arbitrary term $\mathbf{t}$ let $\mathbf{t}_\mathfrak{A}$, resp. $\mathbf{t}_\mathfrak{B}$, denote the term operations induced by $\mathbf{t}$ in $\mathfrak{A}$, resp. $\mathfrak{B}$. We claim that for arbitrary term $\mathbf{s}$, if $\mathbf{s}_\mathfrak{B}$ is a projection, then $\mathbf{s}_\mathfrak{A}$ is also a projection. Indeed, suppose $\mathbf{s}_\mathfrak{A}$ is not a projection. Then by the minimality of $\mathcal{T}(\mathfrak{A})$ we get that $(A; \mathbf{s}_\mathfrak{A})$ is term equivalent to $\mathfrak{A}$. Hence $(B; \mathbf{s}_\mathfrak{B})$ must be term equivalent to $\mathfrak{B}$, implying that $\mathbf{s}_\mathfrak{B}$ is not a projection. Clearly, if $\mathbf{s}_\mathfrak{B}$ is an i-th projection, then $\mathbf{s}_\mathfrak{A}$ is also an i-th projection. Thus, for arbitrary term $\mathbf{p}$ inducing $x-y+z$ in $\mathfrak{B}$, the identities

$$\mathbf{p}(x,y,y)=x=\mathbf{p}(y,y,x), \tag{4}$$

$$\mathbf{p}(\mathbf{p}(z,y,x),z,y)=x, \tag{5}$$

$$\mathbf{p}(\mathbf{p}(\mathbf{p}(x,y,z),z,u),u,y)=x, \tag{6}$$

$$\underbrace{\mathbf{p}(\mathbf{p}(\ldots(\mathbf{p}}_{q\text{ times}}(x,y,z),y,z)\ldots),y,z)=x, \tag{7}$$

which obviously hold in $\mathfrak{B}$, are satisfied in $\mathfrak{A}$ as well. By (4), $\mathbf{p}_\mathfrak{A}$ is a Mal'tsev operation. Identifying the variables u, y in (6) we get the identity

$$\mathbf{p}(\mathbf{p}(x,y,z),z,y)=x,$$

which shows that (5) and (6) are equivalent to

$$\mathbf{p}(z,y,x)=\mathbf{p}(x,y,z) \quad \text{and} \quad \mathbf{p}(\mathbf{p}(x,y,z),z,u)=\mathbf{p}(x,y,u),$$

respectively. These identites imply that for arbitrary element $0\in A$ the operations

$$x+y=\mathbf{p}_\mathfrak{A}(x,0,y) \quad \text{and} \quad -x=\mathbf{p}_\mathfrak{A}(0,x,0) \quad (x, y\in A)$$

define an Abelian group $\mathbf{A}=(A;\ +,\ -,0)$, and $\mathbf{p}_{\mathfrak{A}}(x, y, z)=x-y+z$ (see Proposition 2.2 in [13]). Now the identity (7) ensures that $\mathbf{A}$ is an elementary Abelian q-group. Since $\mathscr{T}(\mathfrak{A})$ is a minimal clone, $\mathfrak{A}$ is term equivalent to the algebra $(A;\mathbf{p}_{\mathfrak{A}})= =(A; x-y+z)$, as required.

Proof of Theorem 3.1. The sufficiency is well known. To prove the necessity consider a finite algebra $\mathfrak{A}=(A; p)$ with minimal clone, where p is a Mal'tsev operation. Let $\mathfrak{B}=(B; p)$ be a minimal nonsingleton subalgebra of $\mathfrak{A}$. Clearly, p is a Mal'tsev operation on B, too, and $\mathscr{T}(\mathfrak{B})$ is a minimal clone. Furthermore, $\mathfrak{B}$ is a plain idempotent algebra. If $|B|\geqq 3$, then by Lemma 3.3 $\mathfrak{B}$ is term equivalent to $(B; x-y+z)$ for some cyclic group $\mathbf{B}=(B;\ +)$ of prime order. Using that p is a Mal'tsev operation, one can easily verify that this is true also when $|B|=2$. (Alternatively, we can draw the same conclusion for $\mathfrak{B}$ by applying R. McKenzie's Theorem [3] stating that every finite plain Mal'tsev algebra is either quasi-primal or affine with respect to an elementary Abelian group.) Thus, by Lemma 3.4, $(A; p)$ is term equivalent to the algebra $(A; x-y+z)$ for some elementary Abelian group $(A;\ +)$. As $x-y+z$ is the unique Mal'tsev operation in the clone of $(A; x-y+z)$, the operation p coincides with $x-y+z$.

Proof of Theorem 3.2. The necessity is obvious. Conversely, suppose that $\mathfrak{A}=(A;\ \cdot)$ is a finite idempotent groupoid with minimal clone such that $\mathfrak{A}$ has a minimal nonsingleton subgroupoid $\mathfrak{B}=(B;\ \cdot)$ with $|B|\geqq 3$. Clearly, $\mathscr{T}(\mathfrak{B})$ is nontrivial, therefore it is a minimal clone. Furthermore, $\mathfrak{B}$ is a plain idempotent algebra. In the same way as in the previous proof, Lemmas 3.3 and 3.4 yield that $\mathfrak{A}$ is term equivalent to the algebra $(A; x-y+z)$ for some elementary Abelian q-group (q prime). Obviously, $q=|B|$ $(\geqq 3)$, which completes the proof.

References

[1] D. M. Clark, P. H. Krauss, Para primal algebras, *Algebra Universalis*, **6** (1976), 165—192.
[2] P. H. Krauss, On quasi primal algebras, *Math. Z.*, **134** (1973), 85—89.
[3] R. McKenzie, On minimal, locally finite varieties with permuting congruence relations, Preprint, 1976.
[4] P. P. Pálfy, personal communication.
[5] A. F. Pixley, Functionally complete algebras generating distributive and permutable classes, *Math. Z.*, **114** (1970), 361—372.
[6] A. F. Pixley, The ternary discriminator function in universal algebra, *Math. Ann.*, **191** (1971), 167—180.
[7] J. Płonka, *R*-prime idempotent reducts of groups, *Arch. Math. (Basel)*, **24** (1973), 129—132.
[8] E. L. Post, *The two-valued iterative systems of mathematical logic*, Ann. Math. Studies 5, Princeton Univ. Press (Princeton, 1941).

[9] I. G. Rosenberg, La structure des fonctions de plusieurs variables sur un ensemble fini, *C. R. Acad. Sci. Paris, Ser. A. B.*, **260** (1965), 3817—3819.

[10] I. G. Rosenberg, Minimal clones I: The five types, in: *Lectures in Universal Algebra* (Proc. Conf. Szeged, 1983), Colloq. Math. Soc. J. Bolyai, vol. 43, North-Holland (Amsterdam, 1986); pp. 405—427.

[11] I. G. Rosenberg, L. Szabó, Local completeness. I, *Algebra Universalis*, **18** (1984), 308—326.

[12] L. Szabó, Interpolation in idempotent algebras, in: *Lectures in Universal Algebra* (Proc. Conf. Szeged, 1983), Colloq. Math. Soc. J. Bolyai, vol. 43, North-Holland (Amsterdam, 1986); pp. 497—507.

[13] Á. Szendrei, *Clones in universal algebra*, Séminaire de Mathématiques Supérieures, vol. 99, Les Presses de l'Université de Montréal (Montréal, 1986).

[14] Á. Szendrei, Locally para-primal algebras, in: *Contributions to General Algebra 4* (Proc. Conf. Salzburg, 1986), to appear.

BOLYAI INSTITUTE
ARADI VÉRTANÚK TERE 1
6720 SZEGED, HUNGARY

Bibliographie

V. I. Arnold, Catastrophe Theory (Second, Revised and Expanded Edition), XIII+108 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

The strenuous interest in catastrophe theory, a mathematical tool of the description of jump transitions, discontinuities and sudden qualitative changes in the evolution of systems, has been continuous since its early development. The first edition of this excellent booklet has helped us form a judgement of the features and limitations of this theory and its applications, clear up the mysticism involved.

This first edition of the book was reviewed in these *Acta*, 47 (1984), 492—493. On the changes let us cite from the author's preface to this second English edition: "The present, most complete edition differs from the 1983 Springer edition at many points. A new chapter on Riemann surfaces, vanishing cycles and monodromy has been included from the second Russian edition, the abundant misprints of the first edition have been corrected and some recent new results are described (results on normal forms for the singular points of implicit differential equations and for slow motions in the theory of relaxation oscillations, results on boundary singularities and imperfect bifurcations, results on the geometrical meaning of the caustic of the exceptional group F_4 and on applications of the symmetry group H_4 of the 600-hedron in optimal control or calculus of variations problems)."

It is worth emphasizing that the new edition is a new translation from the Russian, which follows the standard mathematical terminology more than the first one.

The author wrote this book for a general public having minimal mathematical background, in order to clarify a new branch of mathematics which stands in the limelight. But mathematicians also have learnt a lot from this booklet, namely, the essence of the catastrophe theory. In addition, one can learn from it how to throw light upon a field of mathematics in an enjoyable formula-free way.

L. Hatvani (Szeged)

Banach Spaces. Proceedings, Missouri 1984. Edited by N. Kalton and E. Saab (Lecture Notes in Mathematics, 1166), VI+199 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

This book contains 22 papers presented at the conference "Factorization of linear operators and geometry of Banach spaces" held at the University of Missouri from 24 June to 29 June, 1984. The main themes of the papers are the weak topology, projections, the Radon—Nikodym theorem, Gateaux differentiability, Lie algebra of a Banach space and factorization of operators.

L. Gehér (Szeged)

T. S. Blyth—E. F. Robertson, Algebra through practice. A collection of problems in algebra with solutions, Books 1, 2 & 3, X+97+99+95 pages; Books 4, 5 & 6, X+104+101+100 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1984, 1985.

There are very few problem books in abstract algebra. This series of books meets a long-felt need in teaching and studying algebra, by providing about 640 exercises with complete solutions and almost 150 problems without solutions for those intending to test their proficiency. By means of these exercises the reader not only can practice himself in proving statements but also meets a number of concrete examples of algebraic structures. As the authors say they "have attempted, mainly with the average student in mind, to produce a varied selection of exercises while incorporating a few of a more challenging nature".

The series consists of six books. The exercises cover the most classical parts of algebra: sets, relations and mappings; linear algebra (matrices, vector spaces, linear mappings, Jordan forms, duality); group theory (subgroups, factor groups, automorphisms, Sylow theory, series, presentations), ring theory (ideals, divisibility in integral domains, unique factorization), field theory (extensions, Galois theory) and module theory (exact sequences, diagrams, chain conditions, Jordan—Hölder theorem, free and projective modules). Perhaps Galois theory receives less emphasis than it ought to.

At the beginning of each chapter the notions and results the reader is supposed to be familiar with are summarized. For the convenience of the reader, a list of widely used textbooks is included which the reader may consult for background material. It is indicated which chapters are most relevant to the chapters of the present books.

Mária B. Szendrei (Szeged)

B. Booss and D. D. Bleecker, Topology and Analysis. The Atiyah—Singer Index Formula and Gauge-Theoretic Physics, XVI+451 pages, Universitext, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985.

This text-book offers the interested reader a clear, easily accessible introduction to the Atiyah—Singer index formula, "one of the deepest and hardest results in mathematics" according to Hirzebruch, and also provides a survey on its applications of extremely wide range. Only a minimal amount of prerequisites is needed to read the book: some basic linear algebra and Hilbert space theory, elements from the theory of ordinary differential equations and the notion of a differentiable manifold.

The first three parts of the text give the translation (by D. Bleecker and A. Mader) of the original German edition (by B. Booss): Topologie und Analysis, Eine Einführung in die Atiyah—Singer—Indexformel, Springer-Verlag, 1977. The main points of Part I "Operators with index" deal with the properties of the index of Fredholm operators (e.g., the homotopy invariance of the index is established) and with the structure of the space of Fredholm operators on a Hilbert space. Part II "Analysis on manifolds" is devoted to the elements of the theory of partial differential equations, pseudo-differential operators, elliptic differential operators and boundary-value problems. Part III "The Atiyah—Singer index formula" is the heart of the book. After an introduction to K-theory and the index formula in the Euclidean case, here the embedding, the cobordism and the heat equation proofs of the index formula for an elliptic operator on a closed, oriented Riemannian manifold are all explained. Then a survey on applications including, for example, the cohomological formulation of the index formula, the theorem of Riemann—Roch—Hirzebruch and the Lefschetz fixed point formula is presented. Part IV "The index formula and gauge-theoretical physics" is written by D. Bleecker. Here the author first gives an account of the basic concepts in the geometrization of Yang—Mills theory. Then he expounds in detail how to use the index formula to instanton parameter counting in gauge theory and shows that the moduli space of irreducible self-dual connections (instantons) is naturally a manifold under suitable hypotheses.

The book under review gets the reader with a minimal background of knowledge and experience acquainted with one of the central theorems of differential topology and also with some of its "traditional" and most recent applications. So it can be recommended to everyone interested in the index theorem or in its applications, from students to active mathematicians and theoretical physicists.

L. Gy. Fehér (Szeged)

R. Carmona—H. Kesten—J. B. Walsh, École d'Été de Probabilités de Saint-Flour XIV—1984. Édité par P. L. Hennequin (Lecture Notes in Mathematics, 1180), X+439 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

This new Saint-Flour volume has a very high chance to become as much a widely referenced big success as the Saint-Flour IV—1974 volume (Lecture Notes in Mathematics, 480) which contained Fernique's famous 96-page paper" Régularité des trajectoires des fonctions aléatoires gussiennes". Note the (random?) play of the numbers: this was ten years and ten Saint-Fluor volumes, and exactly 700 Lecture Notes ago.

The present volume contains three long survey articles by three very illustrious mathematicians. The first paper is "Random Schrödinger Operators" by René Carmona (1—124 pages). It gives a rigorous development and an up-to-date overview of a relatively new and fascinating field of research, a field motivated by quantum mechanics and theoretical physics in general and belonging to both probability theory and functional analysis. There are many new results proved here for the first time and many open problems discussed. The second one is "Aspects of First Passage Percolation" by Harry Kesten (125—264 pages). This is also a fine introduction to the subject and a state-of-the-art survey of the field since the author's book (Percolation Theory for Mathematicians, Birkhäusler, Boston—Basel—Stuttgart, 1982) appeared, the Russian translation of which has just been published. The third paper, almost a monograph by itself, is "An Introduction to Stochastic Partial Differential Equations" by John B. Walsh. The author emphasizes that "this is an introduction, not a survey". However, it is his own introduction unifying two different kinds of approaches (one suited for noise with nuclear covariance, the other one for white noise) in a "(nearly) real variable setting" and also containing quite a number of new results. Looking through the foregoing lines, my feeling is that the "high chance" noted above is in fact probability one.

Sándor Csörgő (Szeged)

G. D. Crown, Maureen H. Fenrick, R. J. Valenza, Abstract Algebra (Monographs and Textbooks in Pure and Applied Mathematics, 99), vi+403 pages, Marcel Decker, Inc., New York—Basel, 1986.

The authors' aim was to write a self-justifying textbook on abstract algebra, which covers the fundamental concepts at a level appropriate to an upper-division undergraduate or first year graduate course.

To carry out this intention much space is devoted to solid, fundamental examples and correspondingly less space is available for advanced topics. The chapter headings are: Preliminaries (set operations, functions, partitions, equivalence relations, binary operations, integers) Groups, Group Actions and Solvable Groups, Rings, Factorization in Commutative Rings, Algebras, Modules and Vector Spaces, Field Extensions, Galois Theory. There are two short appendices on Zorn's Lemma, and categories and functors.

Each chapter is ended with a set of exercises which fall into three categories: instantiations of propositions and definitions (which help the reader in deeper understanding of the subject), routine combinatorial drills (the traditional mainstay of this type of text) and extended sequential exercises developing important supplementary topics.

This well-selected material will serve as a good textbook for both students and teachers.

Lajos Klukovits (Szeged)

L. Devroye, Lecture Notes on Bucket Algorithms (Progress in Computer Science, No. 6) 146 pages, Birkhäuser, Boston—Basel—Stuttgart, 1986.

At bucket algorithms data are partitioned into groups according to membership in equal-sized d-dimensional hyperrectangles, called buckets. In this book the connection between the expected time of various bucket algorithms and the distribution of the data is investigated. A variety of probability-theoretical techniques for analyzing various random variables (the average search time, the time needed for sorting, the worst-case search time etc.) related to the bucket structure is given. This is done in a very nice style: the author starts slowly on standard problems (one-dimensional sorting and searching) and moves on to multidimensional applications, in the areas of computational geometry, operations research and pattern recognition.

The book is recommended for people who are interested in computer algorithms.

J. Csirik (Szeged)

H. J. Eichler—P. Günter—D. W. Pohl, Laser Induced Dynamic Gratings (Springer Series in Optical Sciences, 50), XI+256 pages, Springer Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

The classical principle of linear superposition of electromagnetic fields does not hold in a medium which responds nonlinearly to the external perturbation. When two laser beams are arranged to interfere in a nonlinear medium they produce a transient periodic structure, a nonlinear diffraction grating. If a third wave is falling on this structure, in certain cases this wave will be reflected with reversed phase, i.e. the grating acts as a time reversal operator on the electromagnetic field. The effect has many important applications: restoration of distorted optical beams, optical data storing and processing, etc. Beyond the mathematical theory of dynamic gratings a detailed description of materials and practical arrangements are also presented in this book.

M. G. Benedict (Szeged)

Goodness-of-Fit Techniques. Edited by R. B. D'Agostino and M. A. Stephens (Statistics: Textbooks and Monographs, 68), XVIII+560 pages. Marcel Dekker, Inc., New York—Basel, 1986.

The editors write in their preface that when several of the nine authors first decided writing this book in 1976 they asked Egon S. Pearson, to whom the book is dedicated, if he would join them. "He declined and stated his view that the time was not yet ripe for a book on the subject." Judging from the vast amount of techniques covered, and a rather sizable amount that have been left out, initial look may leave a feeling to the contary: perhaps it was too late to write the first book on the subject.

Following a few-page overview by the editors, an overview of their own work and not of the field, and a chapter by D'Agostino discussing graphical plots of the empirical distribution function (EDF) and related functions of the sample (pages 7—62), a short description of χ^2 tests is given by

D. S. Moore (pages 63—95). This is perhaps the most intelligent chapter in the book. The most sizeable chapter by Stephens follows then (pages 97—193) on tests based on EDF statistics. These include tests for simple goodness of fit in general and for composite normal, exponential, Gumbel, Weibull, Gamma, Logistic, Cauchy, von Mises and some other hypotheses such as symmetry, the statistics themselves falling into one of the three boxes defined by the Kolmogorov—Smirnov, Cramér—von Mises and Anderson—Darling statistics. Chapter 5 (pages 195—234) by Stephens describes correlation or regression tests mainly for the uniform, normal, exponential and Gumbel distributions. Chapter 6 (pages 235—277) gives a clever review of transformation methods by C. P. Quesenberry, while in Chapter 7 (pages 279—329) K. O. Bowman and L. R. Shenton discuss techniques based on the sample skewness and kurtosis. The next three chapters (pages 331—366, 367—419, and 421—459) are on tests for the uniform, normal, and exponential distributious by Stephens, D'Agostino and again Stephens, respectively. The treatment of censored samples is insufficient throughout the book. The special chapter (pages 461—496) by J. R. Michael and W. R. Schucany devoted to this topic is disappointing. The last Chapter 12 (pages 497—522) by G. L. Tietjen on the analysis and detection of outliers is such as a chapter on this problem can be written presently. This is followed by an Appendix (pages 523—549) of selected tables and a not too helpful index.

There are many numerical illustrations involving simulated or real data in each chapter. Also the most needed tables for practical implementation are included in the text where they are first required. There are no proofs in the book. It is for the non-statistician practitioner who is supposed to use these tests.

Any professional statistician will find topics insufficiently treated on the level of importance he/she attaches to them, statements concerning the accuracy of this or that approximation which he/she will disagree with, or recommendations concerning the preference of this or that statistic for a given hypothesis. This is completely unavoidable in case of such a book, and this is perhaps what Egon Pearson had in mind: there are no Neyman—Pearson lemmas to make the picture clear, there are too many open problems and unexplored proposals and techniques, too many personal preferences. (It would be no point, therefore, to list my disagreements, part of which are based on my own personal preferences.) The situation may seem ideal for the researcher. In fact, the editors hope that "this book... will act as a base from which ... many questions can be explored". Considering the complete lack of theory in the book, this is perhaps too much to be hoped for. Be it as it may, Pearson's ripe time will probably always remain in the infinitely distant future.

If not every, but certainly many professional statisticians will miss one or two of their important papers or of their friends' from the reference lists presented at the end of the chapters. (Again, it would be fully needless to use space for my friends' missing-lists. However, I cannot refrain from mentioning one outside of that circle. It is the fine booklet "Omega—Square Tests" by G. V. Martynov, Nauka, Moscow, 1978; MR 80 g: 62028. It contains newly computed tables for all Cramér—von Mises and Anderson—Darling tests. Although it is in Russian, the tables of course use Arabic numbers.)

Was Egon Pearson right or not? He was, and he was not. Were the authors right to write this book? Yes, they were. Is it a good book or not? It is a useful book.

Sándor Csörgő (Szeged)

Wolfgang Hackbusch, Multi-Grid Methods and Applications (Springer Series in Computational Mathematics, 4), XIV+377 pages with 43 figures and 48 tables, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

Is there a gap between "pure" and "applied" mathematics? The public is certainly sure that such a gap exists. Generally "pure mathematics" has the image of being a subject which is not associated with practical problems. Perhaps the root of this widely held view is that "mind stands higher than matter". But the history of sciences proves that mathematics underlies in all kinds of scientific and technological developments. For example the theory of differential equations is in the closest connection with practice from the beginnings.

Boundary value problems arise in many fields of applications (in engineering, in physics, etc.). Therefore the solution of these problems is an important task of mathematics. The investigation is twofold: theoretical and practical, this is indeed one of the topics where theory and application are inseparable. The boundary value problems have been the starting points of some other branches of the theory. Then the development of the theory rendered possible not only to prove the existence (or non-existence) and the qualitative properties of the solutions of several problems but to solve them numerically, too. Although in our days we have faster and faster computers, new efficient numerical methods are required having fast convergence.

A certain amount of time is generally necessary to prove the efficiency of a new method. The multi-grid method was first described in the early sixties. Since the seventies a great number of articles verifies its manysided applicability. The main characteristic feature of this method is its fast convergence. The convergence speed, in contrast with some classical methods, does not deteriorate by refining the discretization.

This book gives a clearly written, up-to-date exposition of the subject including several applications, supplied with exercises and interesting comments. The author not only acquaints the reader with a very efficient method but he gives an overall view of the theme. The results which first appeared in various, sometimes hardly obtainable, journals are now made available by this work in a well-organized form for a wide range of interested people. The book will surely remind some readers of the following sentence of P. Halmos: "Pure mathematics can be practically useful and applied mathematics can be artistically elegant."

L. Pintér (Szeged)

Lars Hörmander, The Analysis of Linear Partial Differential Operators Vol. III. Pseudo-Differential Operators, Vol. IV. Fourier Integral Operators, (Grundlehren der mathematischen Wissenschaften, 274, 275), VIII+525 pages, VII+352 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

The first two volumes expand and update the author's book "Linear Partial Differential Operators" which was published in the same *Grundlehren* series originally in 1963 and since then became a standard reference for matematicians working in partial differential equations. The last decades produced an interesting development on the research fields overviewed there and, on the other hand, elegant new techniques appeared which became dominant recently. The author who contributed greatly to this changed fashion of the theory revised his 1963 monograph. The results are Volumes I and II published in 1983 as number 256 and 257 of the same series. Furthermore he has now added these two volumes which can be considered almost entirely new.

In this spirit Volumes III and IV cover the following main branches and their typical applications: Pseudo-Differential Operators and Fourier Integral Operators, Lagrangian Distributions with the underlying Simplectic Geometry, thus areas which became really fruitful only recently after Calderon's Uniqueness Theorem and the Atiyah—Singer—Bott Index Theorem inspite of an existing long tradition in the literature.

Volumes III and IV contain Chapters XVII—XXX of the complete work.

The introductory Chapter XVII, in contrast with all the later ones, displays second order differential operators treated by relatively classical means because of their independent geometrical importance. Chapters XVIII—XX discuss already the powerful machinery of pseudodifferential operataros. After a short heuristical motivation first the necessary basic tools (totally characteristic operators, Gauss transforms, Weyl calculus) and then the index theory of elliptic pseudo-differential operators on compact manifolds without boundary are reviewed which is followed by the treatment of elliptic boundary problems and closed by a motivating outlook to the existence theory of non-elliptic pseudo-differential operators. Based on index theorems, Chapter XXI treats symplectic geometry, a geometrical background essential for later purposes, which has deep roots in classical mechanics but which is now equally important for pure mathematics. Among others, the classifications of pairs of Lagrangian manifolds and of some other systems of relevant geometric systems of mappings are considered together with the symplectic equivalence of quadratic forms. The main aim of Chapter XXII is to illustrate the effectiveness of the methods based on the perturbation theory of pseudo-differential operators by examples of micro-hypoelliptic operator classes occuring naturally in physics and probability theory. Chapters XXIII—XXIV turn to the strictly hyperbolic Cauchy problems and mixed Dirichlet—Cauchy problems, respectively, applying the technique of energy integrals renewed by the calculus of pseudo-differential operators and some extensions of the material in symplectic geometry. Special attention is paid already at this point to the propagation of singularities of solutions. This latter theme for operators of principal type is the main goal of Chapter XXVI whose inclusion to Volume IV is naturally justified by the completeness of the results obtained.

The beginning of Volume IV is Chapter XXV where the author summarizes new arguments concerning Fourier integral operators which had old heuristical motivations in geometry, wave optics and classical and quantum mechanics but whose more systematic study emerged only after the 1960's. Chapter XXVII is devoted to subelliptic operators. In Chapter XXVIII the study of the Cauchy problem is resumed. Problems and tools suggested by Calderon's Uniqueness Theorem are discussed. Chapter XXIX presents very effective applications of the modern theory of Fourier integrals to the asymptotic behaviour of the eigenvalues and eigenfunctions of elliptic operators of higher order. The work is completed by long range scattering theory in Chapter XXX.

Each chapter begins with an about two pages long summary and ends with very detailed historical and bibliographical notes. The two volumes provide a comprehensive reference list of more than 450 items, and a detailed index and list of symbols aids their use as handbooks.

The style of the books can be charactized by their excellent organization which anables us to obtain a relevant insight into almost the whole of the enormous material of wide research areas.

The four volumes can be classified as professional reading. However, taking into consideration the range of direct and indirect applicability of the described results and methods they can be suggested as an indispensable collection of handbooks for all research teams in mathematics and theoretical physics even if their fields of interest are seemingly far from partial differential equations.

L. Stachó (Szeged)

Infinite Dimensional Groups with Applications, Proceedings, Berkeley, 1984. Edited by V. Kac (Mathematical Sciences Research Institute Publications, 4), X+380 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985.

The present book contains the Proceedings of the Conference on Infinite Dimensional Groups held at the Mathematical Sciences Research Institute, Berkeley, May 10—15, 1984. The Conference was concentrated on the following three of the most active directions in the theory of infinite dimensional groups: general Kac—Moody groups, gauge groups and diffeomorphism groups. These are in close connection with physical theories of current interest. Here the key-words are: solitons and instantons, completely integrable systems, Yang—Mills fields and string theory.

The best way to orient oneself is to have a look at the table of contents: 1. M. Adams, T. Ratiu and R. Schmid: The Lie group structure of diffeomorphism groups and invertible Fourier integral operators with applications. 2. E. Date: On Landau—Lifshitz equation and infinite dimensional groups. 3. D. S. Freed: Flat manifolds and infinite dimensional Kähler geometry. 4. R. Goodman: Positive-energy representations of the group of diffeomorphism of the circle. 5. M. A. Guest: Instantons and harmonic maps. 6. Z. Haddad: A Coxeter group approach to Schubert varieties. 7. V. G. Kac: Constructing groups associated to infinite-dimensional Lie algebras. 8. I. Kaplansky and L. J. Santharoubane: Harish—Chandra modules over the Virasoro algebra. 9. S. Kumar: Rational homotopy theory of flag varieties associated to Kac—Moody groups. 10. G. Lusztig: The two-sided cells of the affine Weyl group of type $\tilde{A}_n$. 11. A. Pressley: Loop groups, Grassmannians and KdV equations. 12. P. Slodowy: An adjoint quotient for certain groups attached to Kac—Moody algebras. 13. K. Ueno: Analytic and algebraic aspects of the Kadomtsev—Petviashvili hierarchy from the viewpoint of the universal Grassmann manifold. 14. B. Weisfeiler: Comments on differential invariants. 15. H. Yamada: The Virasoro algebra and the KP hierarchy.

This collection of high level papers gives an up-to-date overview on the present status of the theory of infinite dimensional groups and its applications and so it is recommended to everyone interested in the subject.

L. Gy. Fehér (Szeged)

Gabriel Klambauer, Aspects of Calculus (Undergraduate Texts in Mathematics), X+515 pages. Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

It is interesting that among new textbooks in mathematics there are a great number of books on calculus and real analysis (e.g. 15 among the first 50 volumes of the series Undergraduate Texts in Mathematics). Why? Probably because calculus is needed by almost all sciences, it is the most important tool of applied mathematics, and many other branches of mathematics use analysis. Therefore, it has to be tought on very different levels, for which different texts are necessary. However, a new book can be of interest only if the reader can find topics in it shown by the author from a new point of view, topics or examples which are unusual on the level of an introduction. The reader now gets such a book with two special features. One is that it is written for those who know more than "real" beginners, hence have the techniques of the manipulation of formulas, and at the same time wish to find out the deep roots of the basic concepts of analysis. In many texts such a purpose is often carried out by using a too high level of abstraction, which may result in that the reader cannot connect the new concepts with his/her earlier knowledge. But in this book — and this is the other feature — the author succeds in his purpose using very nice instructive examples and exercises. He includes numerous worked-out examples and concludes every chapter by a lot of exercises, the more difficult of which are accompanied with helpful hints of outlined solutions.

The first chapter is a brilliant geometrical introduction to the logarithmic and exponential functions based upon the specific relation between the hyperbolic segment and the logarithmic function. This approach quickly leads to the evaluation of certain important limits (e.g. $n(b^{1/n}-1)\to \to \ln b$ $(n\to\infty;\ b>1)$ can be obtained easily). The second chapter deals with limits and continuity. Chapters 3 and 4 are concerned with differentiation and its applications. A special section is devoted to the inequality between the arithmetic and geometric means. In Chapter 5 the concept of the Riemann integral is prepared by the quadrature of the parabola (by Archimedes), of the cycloid (by Roberval) and of the function $y=Ax^c$ (by Fermat). The last chapter on infinite series, more than 130 pages in extent, gives the most novelties in the book with its instructive examples and propositions.

To sum up, it is the attractive, interesting and useful examples and exercises that make this text very valuable for students being in transition from elementary calculus to rigorous courses in analysis, and indispensable for those teaching calculus and analysis.

L. Hatvani (Szeged)

W. R. Knorr, The Ancient Tradition of Geometric Problems, ix+441 pages, Birkhäuser, Boston—Basel—Stuttgart, 1986.

In the ancient Greek geometry, to raise a geometric problem was a request for working out a construction of a figure corresponding to a specific description. There were three famous problems: cube-duplication (the Delian-problem), angle-trisection and circle-quadrature. This book is a survey of the efforts made by several ancient Greek mathematicians, e.g., Hippocrates of Chios, Eudoxos, Archytas, Archimedes and Apollonius.

The author emphasizes the mathematical and historical aspects of the ancient writings taking into consideration not only the works of the mathematicians but Greek and Arabic commentaries, too. The final chapter includes aspects of philosophical interest as well.

The chapter headings are the following: Sifting History from Legend, Beginnings and Early Efforts, The Geometers in Plato's Academy, The Generation of Euclid, Archimedes—The Perfect Eudoxean Geometer, Successors of Archimedes in the 3rd Century, Apollonius — Culmination of the Tradition, Appraisal of the Analitic Field in Antiquity.

This valuable and beautiful book, which includes 400 geometric drawings and photographs, is recommended to anyone who wants to get acquainted with the ancient Greek geometry, in particular with its three famous problems.

Lajos Klukovits (Szeged)

Hermann König, Eigenvalue Distribution of Compact Operators, (Operator Theory: Advances and Applications, Vol. 16), 262 pages, Birkhäuser Verlag, Basel—Boston—Stuttgart, 1986.

The classical Riesz theory provides a qualitative description of the spectra of compact operators. Namely, it claims that every non-zero spectrum point of a compact operator on a complex Banach space is an isolated eigenvalue of finite multiplicity. The asymptotic behaviour of the sequence $\{\lambda_n(T)\}_n$ of the eigenvalues of compact Hilbert space operators T was characterized by H. Weyl in terms of the s-numbers of T introduced by Schatten and von Neumann.

This monograph gives an introduction to the theory of eigenvalue distributions of compact operators acting on general complex Banach spaces. The author's contribution to this rapidly developing theory was decisive. The subject is divided into four chapters.

The first chapter contains a brief account of the Hilbert space case. The generalizations of s-numbers to Banach spaces: the approximation-, Gelfand-, Weyl- and entropy-numbers are treated. Furthermore, the definitions and elementary properties of Lorentz spaces and of different operator ideals, among others the class of p-summing operators, are given. The main theorems, the generalized Weyl inequalities on eigenvalue distributions of operators are proved in the second chapter. In the third chapter applications to integral operators with kernels satisfying summability conditions or belonging to Sobolev and Besov spaces are discussed. The last chapter provides further applications to the trace formula and to projection constants.

The book is written in a clear style. It is almost self-contained, only a basic knowledge in functional analysis is needed. This excellent work can be warmly recommended to everyone who wants to get acquainted with this fascinating subject.

L. Kérchy (Szeged)

Serge Lang, A First Course in Calculus (Undergraduate Texts in Mathematics) Fifth Edition XV+624+A99+13 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

Nowadays the teaching of mathematics is much more widespread and varied than it was fifty years ago. Various fields in science and many of the professions demand a certain mathematical knowledge. The main problem is what mathematics should be taught and how? Calculus is certainly one of the topics which we would like almost all students to know to a certain degree.

The subject of any first course in the calculus consists of the basic notions of derivative and integral and some basic techniques and applications accompanying them. A solution of the problem of "how to present them" is what gives the characteristic differences between the books on calculus. This is a teaching problem primarily.

Lang's present book is a source of interesting ideas and brilliant techniques. The main question in this topic is the introduction of the notion of limit. The author's opinion is that any student is ready to accept as intuitively obvious the notions of numbers and limits and their basic properties. Epsilon-delta should be entirely left out of ordinary calculus classes. From the mathematical point of view this is not without danger. But in the reviewer's opinion here the mathematical and methodological difficulties are avoided in a masterly manner.

Let us mention another important feature of the book. It is well-known all over the world that students' facility in speaking and writing is less and less sufficient. The author writes in his Foreword: "I have made great efforts to carry the student verbally, so to say, in using proper mathematical language. It seems to me essential that students be required to write their mathematics paper in full and coherent sentences."

Many of the well-chosen problems and exercises are useful for both students and instructors.

The fact that the book has been reprinted and expanded after over twenty years says all that is needed to say.

L. Pintér (Szeged)

Serge Lang, Math! Encounters with High School Students, XII+138 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985.

It can be often heard that in most mathematical books, especially in school books, topics are treated in an incoherent way. Things are piled up without noticeable reason, technical details are accumulated endlessly, etc. Similar opinion is widely held about the teaching of mathematics, too.

Unfortunately, in many cases these opinions reflect the facts. Therefore it is especially important to see good books and to hear good lectures convincing people (at least some of them) that doing mathematics is a lively and beautiful activity. The production of such works does not depend on the intention of the author solely.

The author of this book experimented a risky undertaking while giving talks to students about 15—16 years old on various deeper problems of mathematics. The author is an experienced teacher (who has written about thirty books), a creative mathematician and these two qualities are inseparable. Therefore he has chances to achieve his aim.

This book contains seven talks (or rather dialogues) given by the author in various high schools in Canada and in France. The titles of the talks give an insight into the questions: What is pi?; Volumes in higher dimension; The volume of the ball; The length of the circle; The area of the sphere; Pythagorean triples; Infinities. Each dialogue is self-contained. The reviewer's favourite is the last one on infinities. The Postscript is a discussion concerning the teaching of mathematics interesting mainly for teachers.

I would recommend this book to students and teachers and I am sure that teachers will find several ideas and patterns helpful in the everyday work.

L. Pintér (Szeged)

Serge Lang, The Beauty of Doing Mathematics, Three Public Dialogues, XI + 127 pages, Springer-Verlag, New York—Berlin—Hedigelberg—Tokyo, 1985.

Almost every Saturday from October to June the Science Museum in Paris welcomes and presents to the public eminent lecturers in all disciplines.

This book consists of three lectures given by Serge Lang. The audience was very diverse, ranging from young students to retired people, from housewives to engineers, but they were people curious enough. How can one convince them that mathematics is quite beautiful?

Serge Lang's aim was to show what pure mathematics is by examples, by doing mathematics with the audience. The first two lectures, Prime Numbers, and Diophantine Equations are in some sense near to a non-mathematical public. For example in the first part of the first lecture the author defined the prime numbers, the twin primes, proved that there are infinitely many prime numbers. He raised the question: Is there an infinite number of twin primes? The activity of the public was shown by several good questions.

The first two lectures are very interesting but the real surprise is the third. In this Professor Lang tries to explain a new geometrical result. I am sure that this lecture is a sensation for almost every reader. Therefore we don't say more. Read it!

L. Pintér (Szeged)

Lyapunov Exponents, Proceedings, Bremen, 1984. Edited by L. Arnold and V. Wihstutz (Lecture Notes in Mathematics, 1186), VI + 374 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

By his famous thesis A. M. Lyapunov founded the stability theory of differential equations. He offered two methods of investigation: the first method was based upon the concept of the so-called exponents of solutions and the second or direct method used an auxiliary function of state variables and the time. Nowadays these exponents and functions bear his name.

The first method started with the following celebrated theorem. If the eigenvalues of A have negative real parts then the zero solution of the system $\dot{x}=Ax+f(t,x)$ $(x\in R^n,\ t\geqq 0,\ |f(t,x)|\leqq \leqq c|x|^{1+\alpha}$ for some $c,\ \alpha>0)$ is exponentially stable. Comparing the solutions of a linear equation $\dot{y}=A(t)y$ with the exponential functions $\exp[\lambda t]$ $(\lambda\in R)$, Lyapunov introduced the exponent of a solution and defined a spectrum for this kind of an equation, too, and was able to generalize the above theorem to the case of varying A.

The first method of Lyapunov has been widely applied and expansively developed further since his pioneering works. This volume, which contains 22 invited papers of the Workshop, gives a good flavour of these kind of results. The editors open the volume by an excellent survey article. In its first part the history and the classical results of the theory are reviewed. In the second part the authors write about the modern areas of the theory and characterize what the papers in the Proceedings contribute to them. The main fields are the following: 1. Products of random matrices and random maps; 2. Linear stochastic systems. Stability theory; 3. Random Schrödinger operators. Wave propagation in random media; 4. Nonlinear stochastic systems. Stochasitc flows on manifolds; 5. Chaos and phase transitions.

The volume is concluded by the complete list of the papers presented at the Workshop and a Subject Index.

L. Hatvani (Szeged)

Saunders MacLane, Mathematics: Form and Function, XI+476 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

One of the main problems of the teaching of mathematics is the fact that students learn relatively much in special branches of mathematics without acquiring a general image of the subject. For example after studying functions, their properties and applications in various fields the question "What is a function?" may be embarassing. Naturally enough, teaching cannot begin with this question but after some experience we must raise it nevertheless. After similar questions finally it is asked: What is the function of mathematics and what is its form? This is the main problem of this book which grew out of lectures given by the author and which is intended as a background for the philosophy of mathematics. Therefore the first task was to make clear what mathematics is. The book contains a survey of some basic parts of mathematics. In the course of the treatment the author tries to answer six questions: What is the origin of mathematics? What is the organization of mathematics? Are the formalisms of mathematics based on or derived from the facts: if not, how are they derived? How does mathematics develop? Is there an absolute standard of rigor and what are the correct foundations of mathematics? The most fundamental is a bundle of questions concerning the philosophy of mathematics: What are the objects of mathematics and where do they exist? What is the nature of mathematical truth? How is it that we can have knowledge of mathematical truth or of mathematical objects?

Let us enumerate the chapter headings: Origins of formal structure; From whole numbers to rational numbers; Geometry; Real numbers; Functions, transformations and groups; Concepts of calculus; Linear algebra; Forms of space; Mechanics; Complex analysis and topology; Sets, logic and categories; The mathematical network.

The book is very valuable for everybody who has some experience in mathematics. Several details are interesting in themselves. Perhaps the reader will not agree with the author in some of the answers but having read the book his/her own view will surely be more well-considered and endowed with new features in many cases. In my opinion this work is a source of important ideas especially for teachers. The style and presentation is fascinating. (Recently I went by train with one of my friends. I took this book along for the trip. My friend had a dip into it, then he grabbed my book and left me to bore myself with his newspapers till the end of the two hours' train ride.)

L. Pintér (Szeged)

P. C. Müller—W. O. Schiehlen, Linear Vibrations, (Mechanics: Dynamical Systems), X+327 pages, Martinus Nijhoff Publishers, Dordrecth—Boston—Lancaster, 1985.

Both in mechanics and engineering vibration analysis has been a central field since the very beginning. The theoretical methods in this field are based upon an exact mathematical description of the considered technical systems. This mathematical description leads to one or more differential equations; therefore, there is a very useful interaction between oscillation, vibration theory and the theory of differential equations. Many problems on differential equations arose in vibration theory and results of the theory of differential equations often open new directions of investigation in vibration theory. To illuminate this interaction it is enough to mention stability theory. The problem of stability of an equilibrium of a vibrating system appeared in mechanics long ago. Since 1892, when the great Russian mathematician and mechanician A. M. Lyapunov introduced the exact mathematical notion of a solution of a differential equation and discovered methods for their investigation, an enormous development can be observed also in stability theory in mechanics and engineering.

In the last decade two challenging phenomena inspired the further development in vibration analysis: increasing demands on precision and the growing use of electronic computers. Improvement in precision can be achieved by more accurate modelling of technical systems, which, first of all, means modelling mechanisms as systems with many degrees of freedom such as multibody systems, finite element systems or continua. The presence of big computers is also a motivation for making use models with more degrees of freedom, which could not be handled numerically earlier.

This book is a theoretical treatment of multi-degree-of-freedom vibrating systems. Part I gives a classification of these systems, which is in accordance with the classification of the modelling equations (time-variant or time-invariant systems; free, self-excited and forced vibrations; conservative-non-conservative systems). In Part II time-invariant vibrating systems are discussed. Besides the classical results vibrations excited by periodic functions are treated, which may display resonance, pseudoresonance or absorption (the last two phenomena can occur only in multi-degree-of-freedom systems). Random vibrations are also investigated by means of covariance analysis and spectral density analysis. Part III is devoted to time-variant systems including a detailed descussion on the parameter-excited vibrations and parameter-excited random vibrations. In the last part the mathematical prerequisites beyond matrix calculus are presented.

The book can be used as a text, too. Each section is accompanied with interrelated exercises and multiplechoice questions.

Applications of the results to some interesting and important models such as motor car, double pendulum, centrifuge, magnetically levitated vehicle, run through the book.

This excellently written and easily readable book is highly recommended to every scientist, engineer and student interested in vibration theory and its mathematical justification.

L. Hatvani (Szeged)

Peter J. Olver, Applications of Lie Groups to Differential Equations, (Graduate Texts in Mathematics, 107), XXVI+497 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

Mathematicians and physicists undoubtedly agree that the theory of continuous groups, now universally known as Lie groups, is one of the most important and powerful tool in modern mathematics and physics, engineering and other mathematically-based sciences. It is enough to mention its applications to such diverse fields as algebraic topology, bifurcation theory, numerical analysis, control theory, classical mechanics, quantum mechanics, relativity, continuum mechanics and so on. Nevertheless, probably only a few scientists know that this theory has its root in differential equations. In the last century the crucial problem of the theory of differential equations was to find more and more techniques to solve particular equations. Different types of equations were discovered which can be integrated such as separable, homogeneous and exact equations. It was Sophus Lie who pointed out that this method could be unified by a general integration procedure based on the invariance of the differential equation under a continuous group of symmetries. Later on the success of this discovery was overshadowed a little by the qualitative theory of differential equations, but nowadays research activity in this direction has been speeding up exceedingly. For example, it was recently pointed out that using the method of generalized symmetries (i.e. the method of including the derivatives of the relevant dependent variables in the transformations), initiated by E. Noether in 1918, one can view certain nonlinear partial differential equations as Korteweg—de Vries equations completely integrable.

This excellent book gives an introduction to the theory of Lie groups and its applications, to such important problems as the determination of symmetry groups, generalized symmetries and conservation laws, integration of ordinary and partial differential equations, reduction in order for systems in Hamiltonian form with emphasis on explicit examples and computations. Each chapter is concluded by further examples and exercises with a very wide range of difficulty (some of the exercises can be considered as research programs for a beginner).

This textbook can be warmly recommended to mathematicians, physicists and students interested in the theory and applications of symmetry methods.

L. Hatvani (Szeged)

Orders and their Applications, Proceedings, Oberwolfach, 1984. Edited by I. Reiner, K. W. Roggenkamp (Lecture Notes in Mathematics, 1142) X+306 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

The volume is the proceedings of a meeting on "Orders and their Applications" held in the Mathematische Forschungsinstitut Oberwolfach, 1984. The conference was organized around the following four topics: Galois module structure, non-abelian class field theory and analytic number theory of orders (6 titles due to J. Brinkhuis, C. J. Bushnell and I. Reiner, M. Desrochers, A. Fröhlich, L. McCulloh, M. Taylor); *K*-theory of orders and connection with algebraic geometry (6 titles due to M. Auslander, J. Brzezinski, J. F. Carlson, W. van der Kallen, R. Oliver, P. Salberger); Applications to group theory and group representations (2 titles; authors: F. de Meyer and R. Mollin, R. Sandling); and Classification of indecomposable lattices (4 titles of G. H. Cliff and A. R. Weiss, E. Dietrich, R. Guralnick, W. Rump). Additionally, the book contains survey articles from the main speakers, most of them mentioned above (H. Lenstra is the only exeption) and a historical note (due to R. Sandling).

The volume is interesting for researchers and experts on these four topics.

P. Ecsedi-Tóth (Budapest)

Probability Theory and Harmonic Analysis, Edited by J.-A. Chao and W. A. Woyczyński (Monographs and Textbooks in Pure and Applied Mathematics, 98), VIII+291 pages, Marcel Dekker, Inc., New York—Basel, 1986.

The volume presents fiteen uniformly high-level contributions by lecturers at the Mini-Conference on Probability and Harmonic Analysis, Cleveland, Ohio, May 10—12, 1983, and by speakers at other seminars of the Probability Consortium of the Western Reserve. The interaction between probability theory and harmonic analysis has been the subject of intensive research in the last decade or so and will obviously remain one for a period of time to come. Those who wish to join this stream will need the present collection without any doubt.

The authors have evidently been asked to call their contributions "chapters" to achieve an effect of unity. However, these "chapters" are fifteen individual expository, survey or research papers covering a range "from martingales, stochastic integrals, and diffusion processes on manifolds, through random walks and harmonic functions on graphs, and random Fourier series, to invariant differential and degenerate elliptic operators, and singular integral transforms". Nevertheless, in spite of the diverse nature of all these topics, there is indeed a kind of an effect of homogenity. Almost everyone working either in harmonic analysis or with probabilities on algebraic structures will find a paper or two in this volume indispensable for him or her. Entirelly subjectively, I single out three of them for special mention: Richard Durrett's review of reversible diffusion processes, Michael Marcus's discussion of infinitely divisible measures on the space of continuous functions induced by random Fourier series and transforms, and the 57-page article of Lajos Takács on the harmonic analysis of Schur algebras and its applications in the theory of probability. This last paper is in fact a prototype of the investigation of the interaction mentioned above.

Sándor Csörgő (Szeged)

Proceedings of the 4th Pannonian Symposium on Mathematical Statistics, Bad Tatzmannsdorf, Austria, 4—10 September, 1983.

Volume A: **Probability and Statistical Decision Theory,** Edited by F. Konecny, J. Mogyoródi and W. Wertz, XI+344 pages.

Volume B: **Mathematical Statistics and Applications,** Edited by W. Grossman, G. Ch. Pflug, I. Vincze and W. Wertz, VIII+321 pages.

Akadémiai Kiadó, Budapest and D. Reidel Publishing Company, Dordrecht, 1985.

The pleasant fate of conference series in their developing period appears to be that they expand and improve, provided of course that the necessary persistence and skill is invested continually into the organizing work including the acquirement of necessary funds to support sufficiently many good participants. Then anything can happen: the series ends abruptly (as was the case with the Berkeley Symposiums), it grows further (as the Vilnius conferences on probability do), or its level stagnates (your example). The Pannonian Symposiums have still a long way to go to be measured to the late Berkeley Symposiums, but they are getting better for sure. (This reviewer attended the first, third and fourth, and reviewed the Proceedings of the 3rd Symposium in these *Acta* 47 (1984), page 513). In fact, the big leap has been the 4th Symposium and the two volumes of its proceedings testify this adequately.

Volume A starts with three invited papers. These are a masterly survey of results on spacings by Paul Deheuvels with a number of new results and indications of the many-sided statistical applications; a comprehensive paper by Ulrich Müller-Funk, Friedrich Pukelsheim and Hermann Witting on locally most powerful unbiased tests for two-sided hypotheses; and an expository note of Pál Révész on the approximation of the Wiener process and its local time with many open problems and conjecturs all arising from the provoking observation that "nobody saw ever a Wiener path". These are followed by twenty-one contributed papers on really diverse problems. Out of these, with upmost subjectivity, we single out for special mention the paper on L_1 regression estimation by Luc Devroye and László Györfi, Norbert Herrndorf's note on invariance principles for strongly mixing sequences, and Detlef Plachky's paper on the converse of the Lehmann—Scheffé's theorem. Even this short list of six papers shows that many of the papers in Volume A could well have appeared under the title of Volume B, and this is completely true *vice versa*.

Volume B proudly boasts with the most enjoyable invited paper of Paul Erdős on probability methods in number theory which is one of his characteristic lists of open problems with many dollar-prizes offered by him for "prove or disprove...". Here we single out, perhaps even more subjectively, the note by Margit Lénárd on L_p spline approximation of stochastic processes and Harald Niederreiter's paper on quasi-Monte Carlo optimization.

Meanwhile the 5th and 6th Symposiums have already taken place. The reviewer was unable to attend these and can thus only hope that the trend continues to be upward.

Sándor Csörgő (Szeged)

P. Rabier, Lectures on Topics in One-parameter Bifurcation Problems, (Tata Institute of Fundamental Research Lectures on Mathematics and Physics, 76), VI+286 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

When models of systems and processes — algebraic, differential, functional equations — depend on a parameter, it frequently happens that there are certain values of the parameter such that small deviations of the parameter from these values cause significant changes in the qualitative behaviour of the solutions of the equation. The goal of bifurcation theory is to identify these bifurcation values of the parameter and to describe the nature of the system near such points.

These notes contain the subject-matter of a series of lectures delivered by the author at the Tata Institute of Fundamental Research Centre, Bangalore, in July and August 1984. The reader gets a good account on some interesting and very new ideas. For exemple, breaking with the traditional exposition of the Lyapunov—Schmidt method the author gives a new algorithm for finding the local zero set of a mapping in certain regular cases. The final chapter introduces a new method in the study of bifurcation problems in the degenerate case. Namely, it is shown how to find the local zero set of an $f\in C^\infty$ real valued function of two variables satisfying $f(0)=0$, $Df(0)=0$, $D^2f(0)\neq 0$ but $\det D^2f(0)=0$ (so that the Morse condition fails).

The book is concluded by some applications and remarks on further developments of the methods.

L. Hatvani (Szeged)

Recursion Theory Week, Proceedings, Oberwolfach, 1984. Edited by H. D. Ebbinghaus, G. H. Müller, G. E. Sacks (Lecture Notes in Mathematics, 1141), X+418 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

The proceedings of a conference on recursion theory that took place in the Mathematisches Forschungsinstitut Oberwolfach from April 15th to April 21st, 1984, include the following titles: Ambos—Spies, K. Generators of the recursively enumerable degrees; Blass, A. Kleene degrees of ultrafilters; Chong, C. T. Recursion theory on strongly Σ_2-inadmissible ordinals; Clote, P. Applications of the low-basis theorem in arithmetic; Dietzfelbinger, M., Maass, W. Strong reducibilities in α- and β-recursion theory; Fejer, P. A., Shore, R. A. Embeddings and extensions of embeddings in the r.e. tt and wtt-degrees; Friedman, Sy. D. An immune partition of the ordinals; Griffor, E. R. An application of π_2-logic to descriptive set theory; Hinman, P. G., Zachos, S. Probabilistic machines, oracles, and quantifiers; Homer, St. Minimal polynomial degrees of nonrecursive sets; Jockus, C. G. Jr. Genericity for recursively enumerable sets; Kechris, A. S. Sets of everywhere singular functions; Kucera, A. Measure, π_1^0-classes and complete extensions of PA; Lerman, M. On the ordering of classes in high/low hierarchies; Nerode, A., Remmel, J. B. Generic objects in recursion theory; Odifreddi, P. The structure of m-degrees; Sacks, G. E. Some open questions in recursion theory; Shinoda, J. Absolute type 2 objects; Simpson, St. G. Recursion theoretic aspects of the dual Ramsey theorem; Slaman, T. A. Reflection and the priority method in E-recursion theory; Wainer, S. S. Subrecursive ordinals.

The volume is recommended to experts and students on advanced level in recursion theory.

P. Ecsedi-Tóth (Budapest)

J. A. Sanders—F. Verhulst, Averaging Methods in Nonlinear Dynamical Systems, (Applied Mathematical Sciences, 59), X+247 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985.

The averaging method is the most important asymptotic method of perturbation theory.

Most differential equations admit neither an exact analytic solution nor a complete qualitative description. However, there are some special classes of equations (linear equations, some autonomous systems, ...) the asymptotic behaviour of whose solutions is known. Perturbation theory is the collection of methods for the study of equations close to equations of these special forms. These latter equations are called unperturbed and their solutions are assumed to be known. Briefly speaking, perturbation theory studies the effect of small changes in the differential equations on the behaviour of solutions.

Model equations often contain a small parameter ε, and the size of the perturbation is characterized by ε. If we investigate the effect of the perturbation over a fixed bounded interval of time independent of ε, we can use the variational equation along the unperturbed solution. However, the investigation of the behaviour of solutions over a large time interval, e.g. of order $1/\varepsilon$, is much more complicated. This is the subject of the so-called asymptotic methods of perturbation theory.

The averaging method gives estimates on the difference between a solution of a nonautonomous equation containing a small parameter and the solution of the autonomous equation obtained by replacing the right-hand side by its integral mean. The method has been used to determine the evolution of planetary orbits under the influence of the mutual perturbation of planets since the time of Lagrange and Laplace, often intuitively. Even nowadays, many physicists and astronomers consider averaging a natural and obvious procedure which need not be justified. However, as is shown in the book by examples and counterexamples, it is important to establish a rigorous approximation theory. The problem of strict justification of the method is still far from being solved.

The first two chapters of the book are of introductory character. The third chapter contains the basic theory of averaging with special emphasis on periodic and almost periodic systems. Chapter 4 deals with the cases when either the original or the averaged equation has an attractor. Chapter 5 is devoted to averaging over spatial variables which allows us to handle systems with slowly varying coefficients. In Chapter 6 the normal forms are considered. Chapter 7 is concerned with Hamiltonian systems in the various resonance cases. Here the method of averaging is used to determine periodic orbits and invariant tori. The book is concluded by many appendices with interesting examples, applications and supplements.

This monograph will be very useful for mathematicians, physicists, astronomers and other users of mathematics interested in qualitative aspects of asymptotic methods.

L. Hatvani (Szeged)

D. H. Sattinger—O. L. Weaver, Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics (Applied Mathematical Sciences, Vol. 61), IX+215 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

The authors undertake to give an exceedingly brief and at the same time consistently constructed introduction to the theory of Lie groups and Lie algebras. They are commanded by the aim that readers interested in applications (first of all analysts or physicists) could go deeply into the subject by connecting it with well-known structures and concepts. However the geometer or algebraist can also appreciate the wide-ranging applications of the theory and can get acquainted with the physical motives behind a number of questions belonging to the topic.

The book has the virtue that, while explaining the results in a homogeneous treatment, the authors bring great care to present their historical development as well. The way in which the authors combine their modern attitude with the explanation of the classical development of the basic results on Lie algebras and Lie groups is interesting for the geometer. A similarly significant feature of the book is its descriptiveness. It shows the essence of the structure of Lie groups by investigating the ones that are significant from the physical and geometrical point of view. This descriptiveness is typical for the investigations of the connection between Lie groups and Lie algebras.

Owing to keeping in view the applications, the representations of Lie algebras play an exceedingly important role. After that the reader learns the general structure of Lie algebras (solvability, nilpotency, Cartan's criteria) and structure of semi-simple algebras (Cartan subalgebras, root space), a whole part deals with the representation theory of semi-simple Lie algebras and the very important spinor representations. The same view also explains the reasons why the authors study the real and complex Lie groups and Lie algebras more comprehensively. The last part completes the material with some important applications such as completely integrable systems, the Kostant—Kirillov symplectic structures and spontaneous symmetry breaking.

There is a number of good exercises at the end of each section.

József Kozma (Szeged)

Winfried Scharlau, Quadratic and Hermitian Forms (Grundlehren der mathematischen Wissenschaften, 270) X+421 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

The purpase of this book is to give a glimpse into the theory of quadratic and Hermitian forms from an essentially algebraic point of view. The material is divided into ten chapters and at the end of the book an Appendix can be found. The first chapter introduces the basic concepts which will be used in the first seven chapters: quadratic forms and symmetric bilinear forms over fields of characteristic unequal to 2. In Chapter two the basic methods and results of the algebraic theory of quadratic forms can be found. In Chapter three a short introduction into the relations between quadratic forms and ordered fields is given. The subjects of the fourth chapter are a deeper investigation of the algebraic theory of quadratic forms and the theory of Pfister forms. Chapters five and six deal with the number-theoretic aspect of the theory of quadratic forms: Instead of the integers and the rational fields more generally an arbitrary algebraic number field and its ring of algebraic integers are considered. Chapter seven is devoted to a general and abstract foundation for the important concepts in connection with bilinear, hermitian and quadratic forms. Chapter eight contains basic results about finite dimensional simple algebras and many interesting connections between the theory of quadratic and hermitian forms on the one hand and the theory of simple algebras and involutions on the other. In Chapter nine the theory of Clifford algebras is developed in an elegant *ad hoc* presentation. Chapter ten returns to hermitian forms and continues the investigations begun in the seventh chapter. The appendix contains some examples.

L. Gehér (Szeged)

Thomas B. A. Senior, Mathematical Methods in Electrical Engineering, VIII+272 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sydney, 1986.

This textbook contains the subject-matter of the one-semester course taught by the author at University of Michigan for students in electrical and computer engineering. At every university such a course has to include the Laplace and Fourier transforms and their applications and some basic knowledge on linear systems. Instructors of a course like this can hardly find a good textbook which is suitable also for undergraduate students not having the basic ideas of complex-variable theory. This book has filled the gap.

Chapter 1 gives a short introduction to complex numbers. Chapter 2 acquaints the reader with the Laplace transform and its applications to differential equations. Chapter 3 deals with the basic concepts and methods of linear-systems theory paying attention equally both to the physical and the mathematical aspects of the subject. The same feature characterizes Chapter 4, which is devoted to Fourier series. Chapter 5 is of more mathematical character, in which the reader gets a good introduction with rigorous theorems and proofs to the analysis of functions of a complex variable. Chapter 6 deals with Fourier transforms. The final Chapter 7 is a short discussion on the connection between Laplace and Fourier transforms. Each chapter contains a number of worked examples and ends with exercises.

This textbook will be very useful for undergraduate students who have a firm background in calculus and differential equations and for their teachers as well.

L. Hatvani (Szeged)

Joseph H. Silverman, The Arithmetic of Elliptic Curves (Graduate Texts in Mathematics), XII+400 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1986.

If a mathematician speaks on mathematics to a non-mathematician public the topic is often taken from number theory, especially from diophantine equations. From the beginnings number theory has formed a great part of mathematics. But its role changed surprisingly in the late forties. Number theory, a characteristic branch of "pure mathematics", has a lot of practical applications in our days. The theory did not loose its attractive features, its influence on mathematics is deeper than ever. An important new result in number theory may arouse the interest of mathematicians all over the world. Consider, for example, the proof of the famous Mordell conjecture by G. Faltings from the last years. As originally formulated the conjecture said that any irreducible polynomial $f(x, y)$ with rational coefficients, having genus greater than or equal to two, has at most a finite number of pairs $x_i, y_i \in Q$ with $f(x_i, y_i)=0$.

The aim of this book is to present an essentially self-contained introduction to the basic arithmetic properties of elliptic curves. Although the author presented approximately half of the material of what he hoped to include, what he wrote is a clear well-organized text offering a good survey of the subject. As prerequisites a first course in algebraic number theory and rudiments of complex analysis are supposed. The reader will find in the first two chapters an introduction to the algebraic geometry of varietes and curves with references. There are numerous interesting exercises at the end of the chapters, some of them are unsolved problems. Similar work in this area has not been published yet which, considering the wast amount of research done in the last decades, is a little curious.

The author says in his Preface: "It is certainly true that some of the deepest results in this subject, such as Mazur's theorem bounding torsion over Q and Faltings' proof of the isogeny conjecture, require many of the resources of modern "SGA-style" algebraic geometry. On the other hand, one needs no machinery at all to write down the equation on an elliptic curve and to do explicit computations with it; and so there are many important theorems, whose proof requires nothing more than cleverness and hard work. Whether your inclination leans toward heavy machinery or imaginative calculations, you will find much that remains to be discovered in the arithmetic theory of elliptic curves. Happy hunting!"

L. Pintér (Szeged)

C. Smoriński, Self-Reference and Modal Logic (Universitext) xii+333 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985.

The volume is devoted to the investigation of self-reference using modal logic. The aim of this investigation is to clarify that "Gödel's Theorem is not artificial; the use of self-reference has not been obsoleted by recursion theory or combinatorics; and self-references is not that mysterious. This monograph reports on the beginnings of a coherent theory of self-reference and incompleteness phenomena, ...".

The book is quite self-contained: Chapter 0 collects almost all the background material required in further chapters. Chapters 1—3, the beginning sections of Chs. 4 and 6 form the core of the material. Chapter 1 develops some of the syntactical tools for Modal Logics (Basic Modal Logic and Provability Logics) while Chapter 2 deals with their model theory in the style of Kripke. Chapter 3 is devoted to questions of arithmetic interpretations of Provability Logics by establishing Solovay's Completeness Theorems stating that Provability Logic is the modal logic of provability in Primitive Recursive Arithmetic. The whole material is generalized to bi-modal logics in Chapter 4. The next chapter deals with Lindenbaum fixed point algebras, and the so called diagonalizable algebra in order to obtain some representation theorems. Chapter 6 treats Rosser sentences. Finally, Chapter 7 is devoted to presenting some applicational oriented material.

The book is clearly written and in a good style. It is recommended to anyone interested in Gödel's Incompleteness results and provability.

P. Ecsedi-Tóth (Budapest)

Frederick H. Soon. Student's Guide to Calculus by J. Marsden and A. Weinstein, vol. I—III, 888 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985—1986.

When Calculus I—III by Marsden and Weinstein came out, we expected a good deal out of it (see the review in these *Acta*, 50 (1986), pages 242—243). Relying on our experiences so far, now we can establish that all of our expectations are fulfilled. The book has been proved especially useful for those students who are willing to choose the best way of learning calculus (and any mathematical subject): attempting to solve problems on their own. The present supplement to the textbook can make this method even more effective.

The sections are of the same structure. Each of them is started with Prerequisites, Prerequisite Quiz, Goals and Study Hints. The prerequisite quiz helps the reader decide if he/she is ready to continue. The goals serve as guidelines during the study of the section emphasizing the most important points. The study hints point out what is worth memorizing, and what is not, from the topic.

Each section provides the detailed solutions to every other odd numbered exercise in the corresponding section of the textbook. Since most of the exercises in the book are written in pairs, the solutions can also be used as a guide to solving the corresponding even numbered exercises.

The sections are accompanied with quizzes, at least one of which is a word problem, for the reader to evaluate his/her mastery of the material. Finally, answers can be found to both the prerequisite and section quizzes.

The chapters are concluded by review sections with questions and answers which may appear on a typical test. The three-hour comprehensive exams, included after every third chapter, help the reader prepare for the midterms and final examinations.

This guide — together with the textbook of Marsden and Weinstein — will be welcomed by students who wish to make their study of calculus easier and enjoyable.

L. Hatvani (Szeged)

Stochastic Analysis and Applications, Proceedings, Swansea, 1983. Edited by A. Truman and D. Williams (Lecture Notes in Mathematics, 1095), III+199 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1984.

This is a collection of thirteen research papers, eight of which were read at the Workshop on Stochastic Analysis, Swansea, 11—15 April, 1983, and the rest are some more recent contributions by the Swansea school itself. As the editors write, "the applications include such diverse topics as stochastic mechanics and the Titius—Bode law (for the distances of the planets from the sun), non-standard Dirichlet forms and polymers, statistical mechanics, quantum stochastic processes, the applications of local time to proving path-wise uniqueness of solutions of stochastic differential equations and its application to excursion theory, Bessel processes and pole-seeking Brownian motion, queues, potential theory and Wiener—Hopf theory". The central theme of investigation is of course Brownian motion from which most of the more general processes required by the above applications take their origin. There are many new results for Brownian motion in this collection. However, beside probabilists, the volume may be of interest to theoretical physicists as well.

Sándor Csörgő (Szeged)

The analysis of concurrent systems, Proceedings, Cambridge, 1983. Edited by B. T. Denvir, W. T. Harwood, M. I. Jackson (Lecture Notes in Computer Science, 207), VII+398 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

The volume contains the Proceedings of a Workshop organised by the Standard Telecommunication Laboratories Limited, held in 1983. The four tutorials give expositions of different approaches to the analysis and description of concurrent systems. They are the well known algebric, net theoretic, temporal logic and axiomatic approaches discussed by prominent authors in these topics. However, the most interesting part of the book is the set of ten problems on concurrency and their solutions. The problems are briefly documented and the various solutions of the participants of the Workshop are described in detail. The problems having both theoretical and practical interest are: two-way channel, network service, firing squad, railway, array processor, packet network, parallel reduction of function combinators, mixing synchronous and asynchronous input, cash-point service and matrix switch. Each problem has more solutions based on different theoretical backgrounds due to different authors.

Á. Makay (Szeged)

The book of L**,** Edited by G. Rozenberg and A. Salomaa, XV+471 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1986.

The book is dedicated to Aristid Lindenmayer who introduced language-theoretic models in biology referred to as L systems. It contains about 40 articles showing a continuous interest in the topic. Most of them are up-to-date research papers concerning different classes of L systems (e.g. 0L, D0L, DT0L) from formal language theoretical point of view.

"A 0L scheme is a pair (X, σ) with X a finite alphabet and σ a finite substitution of X into the free monoid X^*. It is deterministic (a D0L scheme) if $\sigma(a)$ is a singleton set for each $a \in X$, and in this case σ can be considered an endomorphism of X^*. ... A 0L system is a triple (X, σ, ω) such that (X, σ) is a 0L scheme and $\omega \in X^*$ is the axiom of the system. For a 0L system $G=(X, \sigma, \omega)$ one considers the languages

$$L_i(G) = \begin{cases} \{\omega\} & \text{if } i = 0 \\ \{\sigma^i(\omega)\} & \text{if } i > 0 \end{cases}$$

The language of G is the set $L(G) = \bigcup_{i=0}^{\infty} L_i(G)$." (H. Jürgensen, D. E. Metthews)

People interested in applications of the L systems find articles in developmental biology, transplantation and software technology.

Á. Makay (Szeged)

The Influence of Computers and Informatics on Mathematics and its Teaching. ICMI Study Series, IV+155 pages, Cambridge University Press, Cambridge—London—New York—New Rochelle—Melbourne—Sidney, 1986.

Willy-nilly one must be interested in the computer. It enters increasingly in everyday life and of course in mathematics, in research, in the process of applying mathematics as well as in teaching. In mathematics the computer is not only a new tool, it is itself the source of new areas of research. As any new tool it comes with advantages and disadvantages, it can be used well or poorly, it can be overemphasized or ignored.

The plan of the International Commission on Mathematical Instruction (ICMI) is to present a series on topics of mathematical education. The first study deals with the influence of computers on mathematics and its teaching. A discussion document was sent to all national delegates of ICMI. It looks in particular at the three themes: 1. How do computers and informatics influence mathematical ideas, values and the advancement of mathematical science? 2. How can new curricula be designed to meet changing needs and possibilities? 3. How can the use of computers help the teaching of mathematics? Contributions written in response formed the basis of discussions at a symposium held in Strasbourg in 1985.

This book contains the above mentioned report and a selection of papers contributed to the Symposium. Let us enumerate some of them: M. F. Atyah: Mathematics and the Computer Revolution (This is one of the most inspiring lectures in the book having sub-titles: A historical perspective, Mathematics and theoretical computer science, Computers as an aid to mathematical research, The intellectual dangers, Economic dangers, Educational dangers, Conclusion). L. A. Steen: Living with a New Mathematical Species; N. G. de Bruijn: Checking Mathematics with the Aid of a Computer; J. Stern: On the Mathematical Basis of Computer Science; H. Murakami and M. Hata: Mathematical Education in the Computer Age; H. Burkhardt: Computer-aware Curricula: Ideas and Realization.

Perhaps even this short list shows that this book is an interesting collection of different opinions and propositions in a theme standing in the limelight of every mathematician and teacher of mathematics.

L. Pintér (Szeged)

Theoretical Approaches to Turbulence, Edited by D. L. Dwoyer, M. Y. Hussaini and R. G. Voigt, XII+373 pages, Springer-Verlag, New York—Berlin—Heidelberg—Tokyo, 1985.

Observations of turbulence, which is the most natural mode of fluid motion, are very old. One can find references to it already in the Bible; Leonardo da Vinci sketched it in circa 1500. The modern scientific study of turbulence or chaos, which dates from the late 1800s with the work of Osborne Reynolds, can be devided into three distinct movements: the earliest statistical movement is of a strong nondeterministic character, the structural movement is predominantly observational, and the most recent one based upon the results of modern theory of dynamical systems is known as the deterministic movement. In spite of the efforts, the phenomenon of turbulence can be considered as one of the oldest and most difficult open problems of physics.

This book contains the subject-matter of the lectures of the recognized leaders (fluid dynamicists, mathematicians and physicists) in the field of turbulence delivered in a workshop during October 10—12, 1984, which was sponsored by The Institute for Computer Applications in Science and Engineering and NASA Langely Research Center. According to the categories of the theoretical approaches to modelling turbulence, the lectures can be devided into four groups: (1) analytical modelling, (2) physical modelling, (3) phenomenological modelling, (4) numerical modelling.

In the preface the editors give an excellent preparatory summary and evaluation on each article included. The 19 titles of the book are as follows: Dennis M. Bushnell, Turbulence sensitivity and control in wall flows; Gary T. Chapman and Murray Tobak, Observations, theoretical ideas, and modelling of turbulent flows — past, present and future; Joel H. Ferziger, Large eddy simulations: its role in turbulence research; Jackson R. Herring, An introduction and overview of various theoretical approaches to turbulence; Robert H. Kraichnan, Decimated amplitude equations in turbulence dynamics; Marten T. Landhal, Flat-eddy model for coherent structures in boundary layer turbulence; B. E. Launder, Progress and prospects in phenomenological turbulence models; W. D. McComb, Renormalisation group methods applied to the numerical simulation of fluid turbulence; A. Pouquet, Statistical methods in turbulence; William C. Reynolds and Moon J. Lee, The structure of homogeneous turbulence; P. G. Saffman, Vortex dynamics; D. Brian Spalding, Two-fluid models of turbulence; E. A. Spiegel, Chaos and coherent structures in fluid flows; R. Temam, Connection between two classical approaches to turbulence: the conventional theory and the attractors; Hassan Aref, Remarks on prototypes of turbulence, structures in turbulence and the role of chaos; Jean-Pierre Chollet, Subgrid scale modelling and statistical theories in three-dimensional trubulence; John L. Lumley, Strange attractors, coherent structures and statistical approaches; Parviz Moin, A note on the structure of turbulent shear flows; S. B. Pope, Lagrangian modelling for turbulent flows.

Anyone who is not familiar with the history and basic ideas of turbulence and chaos, but wishes to get an excellent overview of them, must read the article of Chapman and Tobak. Of course, experts are also warmly recommended to have this book on their bookshelf.

L. Hatvani (Szeged)

Brian S. Thomson, Real Functions (Lecture Notes in Mathematics, 1170), VII+229 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

The book consists of seven chapters and an Appendix. Chapter one introduces a general structure called local system of sets by the aid of which a variety of general notions of limit, continuity, derivative etc. can be formulated. The second chapter gives a review of the classical material on real cluster points and develops some abstract presentation of this material. The purpose of the third chapter is to generalize the elementary notion of continuity, and to introduce the notion of continuity relative to a local system. Chapter four investigates the variation of a function. For each function and local system a measure can be defined which can be used to discover various differentiation properties of the function. Chapter five presents a systematic and detailed investigation of general classes of monotonicity theorems. Chapters six and seven are devoted to describe the relationships among different types of generalized derivatives. The text ends with an Appendix containing a variety of computations directly related to the notion of set porosity.

L. Gehér (Szeged)

Topics in the Theoretical Bases and Applications of Computer Science, Proceedings of the 4th Hungarian Computer Science Conference, Győr, Hungary, July 8—10, 1985. Edited by M. Arató, I. Kátai and L. Varga, X+514 pages, Akadémia Kiadó, Budapest, 1986.

The volume contains a selection of papers presented at the 4th Hungarian Computer Science Conference. The subject of the conference included various topics ranging from theoretical fields to practically motivated computer applications: formal languages, automata theory, Petri nets, program semantics, models of computation, mathematical algorithms, databases and information retrieval systems, distributed systems, expert systems and artificial intelligence. The following is a list of the invited papers: A. Salomaa: Grammar forms: A unifying device in language theory, W. Brauer and D. Taubner: Petri nets and CPS, R. Albrecht: Formal principles of computer architecture, F. Hossfeld: Parallel algorithms — beyond vectorization, Dj. Babayev and R. Babayev: Generating 0—1 integer programming test problems, Y. Matijasevich: A posteriori version of interval analysis.

Z. Ésik (Szeged)

Andrzej Trautman: Differential Geometry for Physicists *(Stony Brook Lectures).* Monographs and Textbooks in Physical Science, V+145 pages, Bibliopolis, Napoli, 1984.

Differential geometric methods are increasingly applied in modern physics, in particular in relativity theory and high-energy physics. Physicists may, however, have difficulty in reading the available (otherwise excellent) textbooks written by (and for) mathematicians. This gap is filled by Trautman's Stony Brook Lectures. Professor Trautman, a recognized authority who has made important contributions to the field, provides students and researchers with a comprehensive, physics-motivated introduction to the theory of differential manifolds, Lie groups and fibre bundle theory. He explains the use of these structures for gauge fields. The theory of characteristic classes and non-trivial fibre bundles is illustrated on the examples of monopoles and instantons. The well-written and nicely-printed book may be used for a one-semester introductory course for physics students.

P. A. Horváthy (Dublin)

S. M. Ulam, Science, Computers and People, From the Tree of Mathematics, edited by M. C. Reynolds and Gian-Carlo Rota, XXI+264 pages, Birkhäuser, Boston—Basel—Stuttgart, 1986.

This is a collection of 23 essays (originally published between 1946 and 1982) written by the famous Polish born mathematician Stanisław Ulam, whose influence on the development of mathematics and, in particular, the application of mathematics in unconventional areas can hardly be overestimated.

According his own view of Ulam — as we can read in a preface written by Martin Gardner — "I am the type that likes to start new things rather than improve or elaborate, ...". He wrote that "I cannot claim that I know much of the technical matherial of mathematics. What I may have is a feeling for the gist, or maybe only the gist of the gist".

In these sentences Ulam was too modest, he knew much about the technical side of mathematics as well, but in his way of seeking the gist he was able to open several new roads which often led to new branches of mathematics, e.g., the theory of cellular atomata (he proposed it to von Neumann), using the Monte-Carlo method in areas different from probability theory, and nonlinear-processes.

Most of the essays are dealing with physical problems (i.e., Ideas of Space-Time, Thermonuclear Devices), computational problems (i.e., A First Look at Computings, Computers in Mathematics, Computations in Parallel), problems on patterns of growth of figures and biological applications.

We can read three very interesting essays on John von Neumann and his work. Probably this is the unique source where these three masterpieces appeared in one volume. There are also four shorter writings on other eminent scientists (Gamow, Smoluchowski, Kuratowski and Banach).

Reading these essays we can enjoy the brilliant writting style of Ulam, which is a mix of crystal clear prose, subtle humor, and graceful phrasing, therefore this volume — which has three introductory chapters written by M. Gardner, Gian-Carlo Rota and Ulam's wife Francoise Ulam — can be warmly recommended to the whole mathematical community.

Lajos Klukovits (Szeged)

Jan-Cees van der Meer, The Hamiltonian Hopf Bifurcation, (Lecture Notes in Mathematics, 1160), VI+115 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

In these notes the author expounds his new theory that gives a complete description of the bifurcation of periodic solutions for the generic case of the Hamiltonian Hopf bifurcation. In fact, he originally was interested in a particular bifurcation of periodic solutions in the restricted problem of three bodies at the equilibrium L_4. In this problem a particle P of zero mass moves subject to the attraction of two other bodies P_1, P_2 of positive mass rotating in circles about their centre of gravity. Euler described the first three equilibrium points of the particle lying on the line of P_1 and P_2 (E_1, E_2, E_3). Later on, Lagrange found two further equilibria which form an equilateral triangle with P_1 and P_2 (L_4, L_5). If we are interested in the motion near L_4 then we have a Hamiltonian system with a so-called nonsemisimple $1: -1$ resonance. Because of the special properties of this resonance the existing methods had to be reformulated in order to deal with the specific nature of the problem. Applying the normal form theory and some ideas of Weinstein and Moser, the author has succeeded in giving a complete description of the behaviour of periodic solutions of short period in the bifurcation as the family of systems passes through the resonance. Such a bifurcation appears in the restricted problem of three bodies at L_4 when the mass parameter passes through the critical volue of Routh.

These well-written lecture notes must be read by every mathematician, physicist and astronomer interested in perturbation theory of Hamiltonian systems, celestial mechanics and, especially, in the three body problem.

L. Hatvani (Szeged)

Robert L. Vaught, Set Theory. An Introduction, X+141 pages, Birkhäuser, Boston—Basel—Stuttgart, 1985.

Vaught's book is intended to serve for a course at the undergraduate level. The author presents the material in the style of the originator of the subject, Georg Cantor. He writes: "For many years, the widely used introductory books on set theory all presented intuitive set theory. For the past two or three decades, the exact opposite has been true: all such books have given axiomatic set theory. But for the student, the trivial and irritating business of fooling around, as he begins to learn set theory with axioms (saying for example that $\{x, y\}$ exists!) discourages him from grasping the main, beautiful facts about infinite unions, cardinals, etc., which should be a joy."

The core of the material is presented in Chapters 1—7. The first five chapters give a good, intuitive introduction to such topics as sets, operations on sets, cardinal numbers, orders and order types, finite sets and number systems (of the integers, rationals and reals). Axioms appear first in the very short Chapter 6 (five pages only!). The next chapter is devoted to the study of well-orderings and the formal definition of cardinals and ordinals (in the manner due to von Neumann), topics (in particular, results on transfinite recursion) which seem to be "more easily grasped working axiomatically than intuitively".

Chapter 8 gives a short, easy discussion of the axiom of regularity. The next chapter presents results in logic which can be used in consistency and independence proofs. The material on logic is out of the scope of set theory, but concludes with formalisations of the ZFC set theory. Chapter 10 gives the relative consistency of the axioms of regularity and infinity following the "inner model" method of von Neumann. Finally, the last chapter returns to pure set theory and provides additional material on the arithmetic of cardinals and ordinals.

The volume is written in a clear and interesting style and is higly recommended to undergraduate students of mathematics as well as of philosophy.

P. Ecsedi-Tóth (Budapest)

Wolfgang Walter, Analysis I. (Grundwissen der Mathematik, 3) XII—385 pages, Springer-Verlag, Berlin—Heidelberg—New York—Tokyo, 1985.

This book contains essentially the material of an introductory analysis course in the first two semesters. The text is divided into three main parts. Part A is introductory. The main purpose of this is to give the notion of real numbers, the basic concepts of set theory, the notion of functions and some fundamental facts concerning functions. Part B introduces the notions of convergence and continuity and gives the usual elementary theorems. Part C is devoted to the introduction of the notions of differentiation and the Riemann integral and presents the classical theory.

The book is recommended to students in the first two semesters as a handbook.

L. Gehér (Szeged)

INDEX

ACTA SCIENTIARUM MATHEMATICARUM
SZEGED (HUNGARIA), ARADI VÉRTANÚK TERE 1
On peut s'abonner à l'entreprise de commerce des livres et journaux „Kultúra" (1061 Budapest, I., Fő utca 32)

ISSN 0324-6523 Acta Univ. Szeged
ISSN 0001 6969 Acta Sci. Math.

INDEX: 26024

87-106 — Szegedi Nyomda — F. v.: Surányi Tibor igazgató

Felelős szerkesztő és kiadó: Leindler László
A kézirat a nyomdába érkezett: 1986. április 29.
Megjelenés: 1987. szeptember

Példányszám: 1000. Terjedelem: 29,5 (A/5) ív
Készült monószedéssel, íves magasnyomással,
az MSZ 5601-24 és az MSZ 5602-55 szabvány szerint